From headscarves to donation

Ugur, Z.B.

Publication date:
2013

Document Version
Publisher's PDF, also known as Version of record

From Headscarves to Donation

Three Essays on the Economics of Gender, Health and Happiness

From Headscarves to Donation

Three Essays on the Economics of Gender, Health and Happiness

Proefschrift

ter verkrijging van de graad van doctor aan Tilburg University, op gezag van de rector magnificus, prof. dr. Ph. Eijlander, in het openbaar te verdedigen ten overstaan van een door het college voor promoties aangewezen commissie in Ruth First zaal van de Universiteit op woensdag 8 oktober 2013 om 16.15 uur door

ZEYNEP BURCU UGUR

geboren op 31 oktober 1982 te Adana, Turkije.

Promotiecommissie:

Promotor:	prof. dr. Peter Kooreman
Overige Leden:	prof. dr. Shelly Lundberg
	prof. dr. Jan van Ours
	prof. dr. Matthijs Kalmijn
	prof. dr. Marcel Das

Preface

There are very many similarities between research and diving. In both endeavors, you can only enjoy the beauty if you are willing to spend enough effort of familiarizing yourself with the environment. These pages contain the seashells I found inside the deep ocean of knowledge. Although diving into this deep ocean was not always smooth, when I look back over these years, I definitely see how my mind and my personality have been transformed with this experience. Without any doubt, I enjoyed these four years.

During this process I have benefited from the advice, knowledge, and support of numerous people. Some of them I want to express my gratitude explicitly. First, I would like to thank my supervisor Peter Kooreman for his support and patience during the last four years towards the avalanche of research ideas I spammed him with during my study. His remarks taught me to be careful and calm before concluding on or even starting any topic.

I would like to thank my committee members as well: Shelly Lundberg, Jan van Ours, Matthijs Kalmijn and Marcel Das for their feedback during my pre-defense and for their time to read and review this thesis.

I am also indebted to Joyce Jacobsen and Meltem Daysal. Joyce's enthusiasm taught me important lessons about how to learn from setbacks. Meltem's advices on research as well as on job market were invaluable.

We shared room K311 and very many ideas with Ting Jiang. In the 3rd year of PhD we shared the same room with Maria Alywin. She was very accommodating especially after the birth of my son.

I got invaluable advices and support from Katie Carman, Elina van der Heijden while I was pregnant. Jens and Patricia Prufer also shared their experience how to handle child care and academic life.

I thank to CentER graduate school for providing a stimulating research environment. I am grateful to my colleagues from Department of Economics, from the health-labor research group for discussions about research. Furthermore, there have been very many other Tilburg University faculty members, students and presenters at various seminars who have contributed to my academic knowledge. Unfortunately, naming all of them would not be possible here.

I am deeply thankful to my parents, who have supported me in my crazy endeavors from my primary school. I hope I will be able to show them love and respect they deserve.

Lastly, I would like to express my gratitude to my husband Mehmet Emre and our little son Abdullah. Mehmet Emre is a living example of selflessness for me to learn. With the birth of Abdullah, we started to learn everything from scratch. His curiosity became a source of joy for us. I hope I will learn to be a better wife and a mother in the near future.

Contents

List of Figures

List of Tables

1 Introduction

This thesis consists of three empirical studies in the domain of the economics of gender, health and happiness. I explore a broad range of topics varying from the effects of the use of headscarves on economic outcomes, to organ donation and monetary donation. The second chapter is written with a gender perspective. The third and fourth chapters are both about donation behavior. Yet, each of them can be read independently. Since this thesis covers a wide area of research, I used data from many different countries and cultures such as Turkey, Iran, Syria, Egypt, EU countries and the Netherlands.

In the first chapter, I provide an overview of the studies by discussing the motivations for the studies, providing research questions and summarizing the main findings. The second chapter documents differences in educational attainment, labor market outcomes and childbearing for women by their use of headscarves and investigates the impact of the headscarf ban on female educational attainment, labor force participation and childbearing decisions in Turkey. In Chapter 3, I explore the relationship between presumed consent legislation and willingness to donate one's organs, organ donation card holding, actual organ donation rates and transplantation rates. The last chapter looks at the relationship between pro-social behavior and subjective wellbeing and tries to quantify the happiness effect of donating in the Netherlands.

1.1 The Turkish Headscarf Ban

The majority of females in Turkey wear headscarves. However, since 1997, wearing a headscarf has been banned in tertiary education and public institutions mainly due to the interpretation that the headscarf is not compatible with secularism. Naturally, one can expect to observe the effects of this ban on various outcomes for females in Turkey. This ban could have led to increased gender inequality because females are supposed to wear the headscarf whereas males are not. Therefore, males who share similar ideas with those females who want to wear the scarf can continue higher education whereas females cannot. There is evidence that Turkey is doing poorly in terms of gender inequality. According to the World Economic Forum's assessment of gender inequality, Turkey's was ranked as 105th in 2006 among 115 countries and 121th among 128 countries in 2007 among , behind

Bangladesh, Tunisia, Syria, Bahrain, Algeria and Qatar. According to World Economic Forum (2007), Turkey is also below Saudi Arabia, Egypt, Iran, and Indonesia for the educational attainment sub-category. In economic participation and opportunity category, Syria, Tunisia, Algeria, Qatar, Bangladesh rank higher than Turkey. In other words, Turkey lags behind some Islamic countries in terms of gender-based inequalities.

The headscarf has been a subject of heated and mostly ideological debate in Turkey. Dealing with integrating its immigrant population, the headscarf has been also an agenda item for many European countries. The motivation for this study is to provide objective information on the effect of the ban which hopefully could increase the quality of the debate in Turkey and might give policy makers in Europe some idea if they enact or abolish similar laws. We are also motivated by examining a potential by-product of the headscarf ban; the childbearing of women who use headscarf. A large body of literature has established the link between education attainment, labor force participation of women and their childbearing. If the preference for wearing the headscarf is strong enough (women wearing headscarves refuse not to wear it), the headscarf ban policy might give rise to higher fertility among headscarved women. This is potentially a side-effect that is not taken into consideration when the headscarf ban was enacted in Turkey.

In this study, we study the impact of the headscarf ban on female educational attainment, labor force participation (LFP) and child bearing decisions in Turkey employing two methodologies. Firstly, we analyze national aggregate data using difference in differences (DD) methodology with Turkey as the treated unit and neighboring countries as the control group. Secondly, national aggregate data is analyzed with females as treated and males as the control group. Lastly, we bring suggestive evidence from individual level data from five surveys.

The contribution of this study to the literature is twofold. Although this ban has been in place for a considerably long time, no economic analysis has been conducted so far. Firstly, we attempt to analyze the ban in terms of consequences for educational attainment, labor market and childbearing. Secondly, we are the first to document a large discrepancy in schooling, labor market prospects and childbearing outcomes for women by their use of headscarves.

Ideally, we would need individual level data which has information on headscarf use, educational attainment, employment status and childbearing before and after the ban. Unfortunately, we only have individual level data for headscarf use after the ban. We can still provide suggestive evidence because women who are born after 1980 are fully exposed to the ban, whereas women born before 1976 might not be exposed to the ban fully if they did not repeat any grade. Our analysis is based on the assumption that women's religious preferences do not change over time.

The results from country level analysis using difference in differences methodology suggest that the headscarf ban led to a 27% drop in the female to male ratio for tertiary education students, but when country specific time trends are added, the effect is no longer statistically significant. Similarly, although we find 22% drop in female LFP, when country specific time trends are added, the effect is no longer statistically significant. However, we find 0.27 increase in total fertility rate from country level analysis which includes country specific time trends.

The results from national aggregate data using males as control group also did not report any significant effect on overall female tertiary education indicators. We observe 3% drop urban LFP rate of females and 2% drop in LFP of higher educated females after the introduction of the ban compared to males. Both estimates are statistically significant at 10% level.

Although all individual level data indicate a large educational gap between women wearing headscarves and women not wearing headscarves, we did not detect a significant difference for tertiary educational attainment of women who were fully exposed to the ban (wearing headscarves and born after 1980) compared to women who may not be exposed to the ban (wearing headscarves and born between 1973 and 79) assuming standard progression through school. We also documented a wide gap in employment status of women by their use of the headscarf. Even after controlling for religion-related covariates, the use of headscarf is negatively associated with being employed. Using employment history from NFHS-2008, we find that after the enactment of the headscarf ban, employment probability dropped by 4.8% for women wearing headscarves. This is more pronounced for younger cohort women wearing headscarves. We observe a drop of 5.9% for this group after the enactment of the ban. Moreover, using full fertility, employment, marriage, and migration history from NFHS-2008, we observe an increased childbearing probability for younger cohort women

wearing headscarves by 1.4% after the enactment of the ban, although the coefficient is significant at 10% significance level. We support this effect on fertility by looking at childbearing in the previous five years from the time of NFHS-2003 and 2008. Using NFHS-2003, we do not find statistically significant effect on short term fertility of women who are fully exposed (wearing headscarves and born after 1980). But, we did find statistically significant effect on long-term fertility on women who are fully exposed.

In sum, we did not find any statistically significant effect of the ban on female tertiary educational attainment indicators; whereas we did find some effect at the national and individual level on female labor force participation indicators and fertility. Even if one does not consider the consequences of the ban for this particular group of women, the unintended byproduct of this ban – an increased number of people raised by mothers who prefer to use headscarf – is significant. The main message of this study is that when addressing concerns of secularism, the potential effects of banning headscarves on women's educational attainment, employment opportunities and fertility should be considered.

1.2 Presumed Consent and its Implications for Organ Donation and Transplantation

The chronic shortage of human organs is leading to premature death of many patients. Therefore, identifying factors that have a potential to impact lives of patients with organ failure is important from a policy perspective. One policy tool that could be cleverly set for increasing organ donation rates is the legislative defaults. Currently, there are two legislative regimes; "informed consent" or opt-in in which explicit declaration makes the person a potential organ donor and "presumed consent" or opt-out in which explicit declaration is required for not being a potential donor. In practice, to avoid painful and difficult discussions with families' of the deceased, in some presumed consent and informed consent countries, consent from the family of the deceased is taken which is sometimes called "soft opt out" system.

The organ donation consent rate in the Netherlands, an informed consent country, is 27.5% while that of Belgium, a country with a very similar culture and economic development, but which has presumed consent regime, has an effective consent rate of 98% (Johnson & Goldstein, 2003). These findings have led to a lively discussion in informed consent countries whether to change the system.

For instance, in the Netherlands, the Liberal Democratic Party (D66) proposed to change the defaults for organ donation to presumed consent (D66, 2012).

Our main aim in this study is to examine how institutional setting namely; presumed consent impacts cadaveric donations and kidney transplantations using a panel dataset from the EU-27 countries plus Croatia in the period 2000-2010. Since there is no country which changed legislation in the period we consider, we could not estimate country fixed effects which would treat any time-invariant unobserved country level heterogeneity. In particular, our pooled OLS results for identifying the impact of presumed consent would be biased if presumed consent is legislated in countries where there is higher social acceptance of organ donation. Therefore, in this study we follow a three step approach. We firstly study differences in willingness to donate one's organs in presumed and informed consent countries. If we do not find any statistically significant difference in willingness to donate one's organs in presumed and informed consent countries, there will be less concern for pooled OLS analysis. Secondly, we study differences in registering preferences for organ donation in presumed and informed consent countries by looking at organ donation card holding behavior. For presumed consent to have an impact on organ donation rates, we should observe differences in registering behavior. If people do register their preferences for organ donation in case of a mismatch between their preferences for organ donation and legislative default, then it is unlikely to observe any behavioral effects of presumed consent. The third step which forms our main analysis explores the impact of presumed consent legislation on cadaveric donations and kidney transplantations.

In the first step, using individual level data from the 2002, 2006, 2009 Eurobarometer Surveys, we do not detect any statistically significant relationship between willingness to donate and presumed consent legislation even after controlling for socio-economic background indicators. These findings imply that presumed consent legislation is not necessarily enacted in countries where there is wide social acceptance of organ donation. This is a useful first step which is necessary but not sufficient condition for claiming that the difference in organ donation outcomes between presumed consent countries and informed consent countries is due to presumed consent legislation.

In the second step, we examine organ donation card holding from Eurobarometer 2006 survey. In line with our expectations, we find significantly lower donation card holding among those who are

willing to donate their organs in presumed consent countries. Surprisingly, among people who are not willing to donate, we do not observe higher registration in presumed consent countries. These findings suggests that presumed consent can increase cadaveric donation rates because people who are not willing to donate their organs fail to register their preferences in presumed consent countries.

In our main analysis, using international organ donation registry data, we find that presumed consent countries have 28 to 32% higher cadaveric donation and 27 to 31% higher kidney transplant rates in comparison to informed consent countries after accounting for potential confounding factors.

Although previous studies also found higher cadaveric organ donation rates in the presumed consent countries compared to informed consent countries, there is no consensus about the underlying mechanism. Some researchers (the first group) attribute higher cadaveric organ donation rates to the effect of presumed consent legislation whereas others (the second group) see the presumed consent legislation as an indicator of a country's commitment to organ donation. Mainly, the first group of studies is criticized on the ground that they did not address unobserved heterogeneity adequately. That is, the results could be due to the presumed consent legislation being enacted in countries where there is higher social acceptance of organ donation.

This study contributes to the literature in some important ways. We firstly address potential endogeneity of presumed consent by showing evidence that presumed consent is not necessarily legislated in countries where there is higher social acceptance of organ donation. To address unobserved heterogeneity even better, we group countries according to their geographic, ethnic, cultural, and organ donation related cooperation. We then identify the impact of presumed consent running country group fixed effects models. The results still show higher cadaveric donation rates in presumed consent counties which suggest that accounting for unobserved heterogeneity ultimately appears to have little effect.

Secondly, we address the claim that presumed consent is an indicator of a country's commitment to organ donation rather than a causal mechanism in itself. We show that after taking into account a country's commitment to organ donation proxied by kidney transplant centers as an additional control variable, the coefficient of presumed consent is still statistically significant and it even increases. Thirdly, although according to Eurobarometer, 2009 survey on organ donation, religious reasons,

distrust in the system and scare of manipulation of the human body are three major causes of refusal for organ donation, previous studies have not dealt with trust in the system and religion differences adequately. Abadie and Gay (2005) include religion with a Catholic country indicator which is based on majority of population being Catholic or not. To capture trust in the system, we included corruption perceptions scores from Transparency International. To control for religiosity changes over time, we compiled percentage of population being Roman Catholic and having no religion mainly from International Social Survey Program (ISSP), European Social Survey (ESS), European Values Survey (EVS) and Eurobarometers conducted between 1999/2000 and 2010. Lastly, to the author's knowledge, this study is the first to analyze the impact of presumed consent on kidney transplantation which is more relevant from a policy perspective.

1.3 Warm Glow

Why do people give away their money or time for free? One potential reason for pro-social behavior is that people get psychological benefits from helping others out, the so-called "warm glow" motivation. To put the warm glow into context, we have to make a distinction between pure altruism and impure altruism at this point. Economists describe a person as pure altruist if she only cares for the final situation of the other person regardless of what she personally did for the other person, whereas an impure altruist would enjoy not only the final situation but also enjoys her own altruistic deed. In large economies, the warm-glow motive must dominate at the margin (Ribar & Wilhelm, 2002). The intuition is that the incentive to free ride must be so overwhelming if large numbers of others are collectively providing a substantial amount of charity, the only justification for giving is that donors get some direct benefit from giving.

Initially, Andreoni (1989) hypothesized "warm glow" to explain incomplete crowding out when the government increases contributions to charity. Later on, the economics literature focused on measuring the extent of crowding out to demonstrate the existence of the warm glow. However, to the author's knowledge there is no empirical study which looks at the relationship between donation and subjective well-being to measure warm glow motivation using real donation amounts.

If the warm glow hypothesis is correct, we should observe higher subjective wellbeing after donation among donors. Higher subjective wellbeing after donation is also consistent with both pure and impure altruism.

There is a large psychology literature on the pro-social behavior and happiness relationship. The literature suggests three mechanisms for this relationship. Firstly, engaging in donation could cause increased happiness. It is also possible that shocks to happiness could lead to higher donations. Moreover, personality characteristics could be driving both happiness and donation. To date, evidence on the first mechanism from psychology literature has been largely based on experimental studies with relatively small sample sizes and unrealistic amounts donated. Thus, whether donating actually causes happiness remains partly unanswered.

If warm glow is an important channel driving pro-social behavior, this could have a lot of implications for policy making. Firstly, one can ask whether there should be tax-break advantages for donations since donors are supposed to get utility from donation anyway. Secondly, this information could also be used for promoting donation.

In this study, we examine pro-social behavior in the form of making donations and try to measure the magnitude of "warm-glow" motivation. We test this theoretical argument empirically by using happiness scores as dependent variable from the Longitudinal Internet Studies for the Social Sciences (LISS), subjective health scores from the Giving in the Netherlands panel (GINP) datasets and life satisfaction scores from the single available wave of the GINP in 2006. To account for endogeneity of donation decisions, we exploit variation in different types of solicitation by charities.

Donating an extra Euro and engaging in donation are our variables of interest. We initially discuss fixed effects regression results from the LISS panel, which primarily measure long-term effects. Secondly, we discuss regression results from fixed effects, OLS and IV specifications from the GINP panel respectively.

From the LISS panel, we do not find any statistically significant relationship between an extra Euro donated and happiness scores after accounting for individual fixed effects; whereas we find evidence that engaging in donation is associated with higher happiness scores after taking into account individual fixed effects. From the GINP, we do not find any effect of donating an extra Euro

nor engaging in donation on subjective health using fixed effects specifications. We find a concave relationship between donating an extra Euro and life satisfaction using the OLS method for the single available wave of the GINP in 2006 whereas from the same dataset using the OLS, we do not detect any significant relationship between engaging in donation and life satisfaction. Contrary to intuition, the IV estimates show a negative effect of donating an extra Euro on subjective health. When experience of certain diseases is taken into account, the effect is no longer significant. This suggests that being solicited is negatively correlated with health status. Lastly, the IV estimates from single available wave of the GINP suggests that an extra Euro donation increases life satisfaction significantly. IV estimates of the effect of an extra one Euro donation on self-reported health being negative and IV estimates of the effect of an extra Euro donation on life satisfaction being positive is at odds with each. This suggests that solicitation reflecting selection cannot explain the results that we observe for life-satisfaction.

To put the findings from the IV methodology into context, one can convert the effect of donation into monetary units by comparing it with the coefficient of income. Our findings from the IV strategy suggest that an increase of €1 in donations increased life satisfaction as much as a €104 increase in income from employment. At first, the effect might seem implausible. However, the results show the local average treatment effect (LATE) for individuals who donate an extra Euro because they are solicited and these individuals would not donate this extra Euro if they were not solicited. This group of individuals is not likely to be representative of the Dutch population. Thus, the IV estimates might not reflect the average treatment effect. Therefore, extrapolation is not meaningful.

For discussing the policy implications, the average treatment effect would be more useful. Nevertheless, the main message of this study - donation makes at least some people happier-, could have policy implications for boosting charitable giving under certain assumptions. Assuming rationality, people are thought to make optimal decisions in which they are supposed to take into account the psychological benefits of donating in their utility function. At first, since a large number of people already donated and people can learn the psychological benefits of donation over time, rationality seems to be a valid assumption. Rationality assumption does not leave room for advertising psychological benefits of donation to affect donation amounts. However, evidence reveals that people

overlook the benefits of charitable giving. Dunn, Aknin and Norton (2008) showed that people erroneously thought that personal spending would make them happier than pro-social spending although they found higher happiness levels of randomly assigned pro-social spenders. Frank (2004) discusses the evidence on how people do not spend their money in ways that yield significant and lasting increases in measured satisfaction. Therefore, if rationality assumption does not hold in the context of pro-social behavior, there might be still room for increasing subjective wellbeing by engaging in pro-social behavior. However, for now, whether people rationally expect these psychological benefits is a potential future research topic.

2 Unveiling the Veil: Implications of the Turkish Headscarf Ban

2.1 Introduction

In 1997, there was a sharp policy change in Turkey that could potentially have large negative effects on women's education attainment and employment opportunities: the headscarf, a religious and cultural artifact, was prohibited in universities and public institutions. As a result of this ban, women in universities and women working as civil servants were forced to resign or drop out of their schools if they refused to uncover their heads.

This ban is not a minor policy change if one considers that, according to Carkoglu & Toprak (2006), 63.5% of females in Turkey wear some sort of headscarf. According to Konda's survey, conducted in 2007, 69.6% of women in Turkey wore headscarves. Similarly, A&G's surveys found that the percentage of households in which women did not wear headscarves was 21.5 and 16.6 in 2003 and 2007, respectively. Another indicator that shows the magnitude of the problem is that 24.5% of the respondents in the study of Carkoglu & Toprak (2006) said they would disapprove if their daughter did not wear a headscarf in order to continue her education in a university.[1] Again, 26.1% of the respondents in Konda's 2007 survey reported that they would prefer their daughters to forgo their university education rather than agree not to wear a headscarf. Certainly, these figures reflect that some part of Turkish society puts considerable emphasis on the use of the headscarf.

This ban was not enacted as a result of societal consensus but was implemented as a result of a National Security Council meeting[2] without much discussion in the public before the decision.[3] The Turkish military was concerned that the headscarf is not compatible with secularism. Because of the complex power relations between the government and the military, it was possible that the decisions taken at that particular meeting were applied without any major objection. The focus of this paper is not to examine how this policy came into place, but to explore its implications.

[1] "If you had a daughter wearing headscarf, would you approve of her not using a headscarf in order to continue her university education?"
[2] National Security Council is composed of government representatives, the president and representatives from the military.
[3] There were events before that Security Council meeting at that time which were considered dangerous for the future of the country, mainly by the military. That meeting and those decisions had far-reaching consequences which are beyond the interest of this paper.

The majority of the public opposes the headscarf ban in schools and the public sphere. Research conducted by Carkoglu & Toprak (2006) shows that 67.9% of the public believes that female civil servants should be able to wear the headscarf, if they want to[4]. Moreover, according to Konda (2007), 78% of the respondents are against the headscarf ban in the universities.

Our study is the first to document the relationships between schooling, labor market prospects and childbearing dispersion and the use of a headscarf in Turkey. This study provides policy makers in other countries a clearer understanding of implications of enacting or abolishing similar laws. The headscarf ban is an issue not only in Turkey but also in other countries. For instance, France has enacted a similar law but not in higher education institutions. Teachers wearing headscarves has also become an issue in Germany (Human Rights Watch, 2009).

On the one hand, it might seem natural that this policy change will affect educational attainment of women who prefer to use the headscarf. However, the effect depends on the strength of the individual's preference for the use of the scarf. Wearing a headscarf in itself could also be affected by the ban.

One can also expect that educational restrictions on this large group will be reflected in their labor market outcomes. That is, fewer females would be able to enter higher end of the labor market since they would not be able to get higher education. On the other hand, those women who cannot continue tertiary education might enter into the labor force earlier, in the lower end of the labor market. Therefore, theoretically speaking, the effect of the ban on female labor force participation is ambiguous.

Moreover, a large body of literature has established the link between employment, education and childbearing: lower levels of employment and education lead to higher birthrates. If the preference for wearing the headscarf is strong and leads many women away from work and/or education, this policy might give rise to higher fertility among headscarved women. Ultimately, whether the ban affects behavioral outcomes is an empirical question. Therefore, in this paper, we try to identify the impact of

[4] Also, 70.4 % of respondents would not be disturbed by a female teacher wearing a headscarf in the classroom of their kid and 71.5% would not be disturbed by a female judge wearing a headscarf.

headscarf ban in Turkey on higher educational attainment of women, labor force participation and their childbearing.

Although this ban has been in place for a considerably long time, no economic analysis has been conducted so far. Cindoglu (2010) studied the headscarf ban through in-depth interviews with a focus group of 79 women. But, the group was not representative and the number of observations was small due to the nature of that study. According to Kaynakoglu & Toprak (2004), the percentage of students who could not continue higher education because of the headscarf ban is only 1%. However, there are several flaws in that study. First of all, although this question is about possible reasons for not continuing to higher education, one category of answer is "currently student" (9.8%). It is ambiguous whether this 9.8% are high-school students or university students. In any case, this 9.8% is irrelevant for analyzing the reasons for not being able to transfer to higher education institutions. Moreover, 10.5% of respondents said that their parents would not allow them to continue on to higher education. Similarly, 49.2% of the female students' reason for not continuing secondary education is "my parents did not allow me." There could be many reasons for some parents not letting their daughters to continue higher education. But one strong consideration for conservative parents is the headscarf ban. There are definitely high costs for getting higher education (such as the time and money involved in entrance examinations, expenses for living, accommodation, tuition and other school fees plus opportunity costs etc.) However, among conservative parents, many would be reluctant to invest in their daughter's education if it meant ceasing to wear the headscarf. Lastly, the main topic of Kaynakoglu & Toprak (2004) is not the headscarf ban, but rather the status of women in the labor market, senior management and politics. They only asked one question about reasons for not continuing higher education, which is not sufficient to evaluate the ban. Carkoglu & Toprak (2006) and Konda's (2007) studies did not look into the effects of the ban but rather at how the ban is perceived in society. Therefore, we cannot know from this study how the headscarf ban affects educational attainment, labor force participation and childbearing rate of women.

Ideally, we would need individual level data which has information on headscarf use, educational attainment, employment status and childbearing before and after the ban. Unfortunately, we only have individual level data for headscarf use after the ban. We can still provide suggestive evidence because

women who are born after 1980 are fully exposed to the ban, whereas women born before 1976 might not be exposed to the ban fully if they did not repeat any grade. However, we observe headscarf use status for once at the time of the survey. Unfortunately, we cannot make an analysis how wearing headscarves has changed after the ban since there is no information about the use of headscarves before the ban. Our analysis is based on the assumption that women's religious preferences do not change over time. This assumption is based on evidence from World Values survey. Details of which are discussed in Table B-1 in Appendix B. Moreover, we have descriptive information from a survey conducted by Anar Research Company in 2007 about the strength of preference for the use of headscarves with the ban. According to that survey, 41% of women who wore scarves at the time of the ban continued to wear the scarf, 35% uncovered their heads in places where the ban was enforced, and 20% continued their education by using wigs or hats as an alternative, so that their natural hairs were not visible (Hazar Group, 2007). Furthermore, we did not find any evidence that women who are fully exposed to the ban (born after 1980) use headscarves less often using NFHS-2003 & 2008. Details of which are provided in Table B-2 of Appendix B.

We study the impact of the headscarf ban employing two methodologies. Firstly, we analyze country level aggregate data using difference in differences (DD) methodology with Turkey as the treated unit and neighboring countries as the control group. Secondly, we provide the effect of the ban on females by using males as a control group using national aggregate data. Lastly, we bring suggestive evidence using individual level data from five surveys.

To evaluate the impact of the ban on tertiary education, the female to male ratio in total number of students are studied at an aggregate level, and university or higher degree attainment is studied using individual level data. For analyzing the impact of the ban on labor market, we focus on female labor force participation rate at an aggregate level, and employment status at the individual level. We also examined total fertility rate at an aggregate level and childbearing at the individual level. The data for most of the aggregate level analysis is obtained from World Bank datasets. Our individual level data comes from five surveys which contain information on educational attainment, employment status, childbearing and headscarf use status. These surveys are Konda's survey conducted in 2007, A&G's

surveys conducted in 2003 and 2007, and the National Family and Health Survey (NFHS) conducted in 2003 and 2008. A more detailed explanation of the data is provided in Appendix A.

The results from country level analysis using difference in differences methodology suggest that the headscarf ban led to a 27% drop in the female to male ratio for tertiary education students, but the effect is no longer statistically significant, when country specific time trends are added. Similarly, although we find 22% drop in female LFP, the effect is no longer statistically significant with country specific time trends specification. However, we find 0.27 increase in total fertility rate which includes country specific time trends.

The results from national aggregate data using males as control group also did not report any significant effect on overall female tertiary education indicators whereas urban LFP rate of females and LFP of higher educated females are impacted with the introduction of the ban compared to males. Both estimates are statistically significant at 10% level.

Descriptive statistics from all individual level data indicate a large educational gap between women wearing headscarves and women not wearing headscarves. However, we did not detect a significant difference for tertiary educational attainment of women who were fully exposed to the ban (wearing headscarves and born after 1980) compared to women who may not be exposed to the ban (wearing headscarves and born between 1973 and 79) assuming standard progression through school. We also documented a wide gap in employment status of women by their use of the headscarf. Even after controlling for religion-related covariates, the use of headscarf is negatively associated with being employed. Using employment history from NFHS-2008, we find that after the enactment of the headscarf ban, employment probability dropped by 4.8% and 5.9% drop for all women and younger cohort women wearing headscarves respectively after the enactment of the ban. Moreover, using full fertility, employment, marriage, and migration history from NFHS-2008, we observe an increased childbearing probability for younger cohort women wearing headscarves by 1.4% after the enactment of the ban, although the coefficient is significant at 10% significance level. We support this effect on fertility by looking at childbearing in the previous five years from the time of NFHS-2003 and 2008. Using NFHS-2003, we do not find statistically significant effect on short term fertility of women who

are fully exposed (wearing headscarves and born after 1980). But, we did find statistically significant effect on long-term fertility on women who are fully exposed to the ban.

In short, we did not detect any statistically significant effect of the ban on female tertiary educational attainment indicators; whereas we did find some effect at the national and individual level on female labor force participation indicators and fertility. Even if one does not consider the consequences of the ban for this particular group of women, the unintended byproduct of this ban – an increased number of people raised by mothers who prefer to use headscarf – is significant. The main message of this study is that, the potential effects of banning headscarves on women's educational attainment, employment opportunities and fertility should be considered when addressing concerns of secularism.

2.2 Background of the Headscarf Ban

The first time the headscarf became an issue was in 1964, in Istanbul University's graduation ceremony. A female student wearing a headscarf who graduated with the highest GPA was not allowed to speak to the audience, although traditionally the student with the highest GPA would give a speech in the graduation ceremony (Cindoglu, 2010).

Although there were one or two incidents until 1980s, the use of the headscarf in universities did not become a problem until the 1980s, because there were only a small number of women in higher education. Among them, women who wore a headscarf were even fewer.

The first regulation about the headscarf was put into effect in 1981 by the Ministry of Education (MONE), in the "Dress Codes for schools under supervision of MONE and other Ministries." In the Official Newspaper, "Resmi Gazete" in Turkish, where amendments to laws and regulations are published, the new regulation explicitly mentioned that the dresses of women should be clean, tidy, ironed, hairs should be combed and inside the institution, the head should not be covered (Official Newspaper, 1981). This regulation covers all students in schools under the control of MONE, which effectively means all schools in Turkey.

The dress code for women working in public institutions which contains articles that women's head should be visible was enacted in 1982 (Official Newspaper, 1982). Again in 1982, the dress code

regulation of MONE was amended so that tertiary education institutions were taken out of the dress code regulation (Official Newspaper, 1982). But, with the establishment of the Higher Educational Council (HEC) in 1982, the new "contemporary" dress codes for tertiary education institutions again stated explicitly that "the head should be visible or open and the headscarf should not be used in the institutions" (HEC, 1982). Due to this regulation, some universities prohibited the use of headscarf very strictly, whereas some of them sort of closed their eyes. The prohibition was implemented differently in different universities, rather than being implemented in a uniform manner throughout the country. However, it is hard to find any data for this time period, because university presidents had discretion to apply the ban.

Over time, student protests led the HEC to circulate a memorandum to universities in 1984 that would let tertiary education students to wear the headscarf in a "modern" way (HEC, 1984). In 1987, article 7/h had been added to the Student Discipline Code of the tertiary education institutions by the HEC, which required students to wear so-called "modern" dresses in classrooms, laboratories, clinics and corridors of institutions, and also mentioned that neck and hairs could be clothed with a headscarf. (Official Newspaper, 1987). In 1988, the headscarf became legally free with the enactment of law no 3511 by the Turkish parliament. The president signed the new law and put it into effect (Official Newspaper, 1988). Then, the president applied to the Constitutional Court for the added article about the headscarf, and the Constitutional Court annulled the article that allowed headscarves to be used. Again, in December 1989, HEC amended the Student Discipline Code and the part about the dress code was removed from the Student Discipline Code (Official Newspaper, 1989). This marks the beginning of a relatively free period of wearing headscarves in tertiary education and public institutions.

Turkey entered a new phase with the National Security Council meeting on 28 February 1997. 18 decisions were taken to prevent the breaching of the principle of secularism in the constitution (National Security Council, 1997). Although the headscarf is not explicitly mentioned in any of the decisions, the 13th decision was related to the so-called modern way of dressing[5]. Headscarves have

[5] Another decision about education from that meeting was that compulsory primary education has been increased from five years to eight years.

been interpreted as against secularism and a modern way of dressing. Therefore, the use of headscarves in the universities was prohibited very strictly nationwide right after the decisions. In the meantime, there were no laws enacted which explicitly outlawed the headscarf[6]. The ban was enforced solely by the National Security Council decisions, which were taken in one meeting without much discussion in the public before the decision. In 2008, there were lawsuits against the headscarf ban in universities; the Supreme Court decided that the ban should be enforced on the ground that the use of headscarves violates the principle of secularism in the constitution.

When it comes to the level of enforcement, particularly after 28th of February 1997, female students wearing a headscarf were prohibited to enter university campuses. Together with physical interference in case of attempts, there were also psychological pressures. For instance, in Istanbul University, so called "persuasion rooms" were formed in order to convince students with headscarves already admitted to the universities by passing the University Entrance Examination not to wear it anymore (Cindoglu, 2010). Female students wearing headscarves organized protests, some of which resulted in police forces arresting protestors. For civil servants, according to AK-DER (2010), between 1998 and 2002, 5,000 women who wore a headscarf were sacked and 10,000 have been forced to quit, because going to work with a headscarf was considered as misbehavior or disobedience.

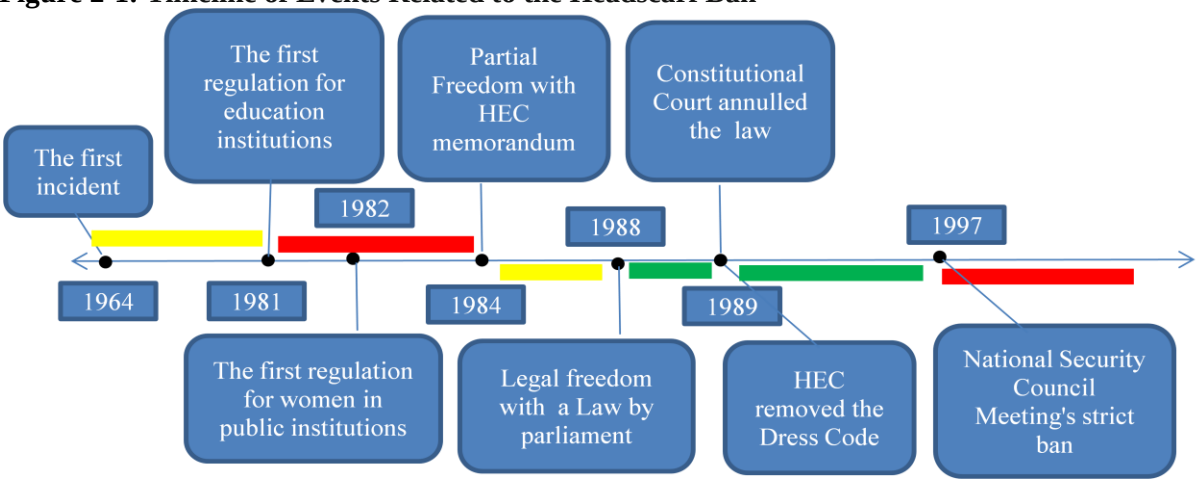
In 2006, "Civil Servants Amnesty" was put into effect which grants civil servants who conducted misbehavior a release from punishment. (Official Newspaper, 2006). Therefore, those headscarved women who had been expelled from their jobs were given another chance to get back their jobs if they would agree not to wear the headscarf anymore.

Moreover, apart from an impact on the labor market through educational restrictions, there have been also more direct limitations in employment opportunities in the public sector for headscarved women. From 2000 onwards, women who are candidates for being civil servants have been obliged to enter placement examinations "with uncovered heads." Secondly, in professions requiring practitioners to be registered to professional organizations, such as doctors, pharmacists, dentists, lawyers, and notaries, the professional Chambers and Unions have issued circulars outlawing the

[6] For the sake of brevity, we refer to these regulatory changes as "the impact of the ban" and the scarf for headscarves.

headscarf. Implementation of these policies also restricts private employment possibilities for headscarved women whenever there is a contact with a public institution. Cindoglu (2010) discusses the propagation of the ban to the private sector through in-depth interviews conducted with women wearing a headscarf and their labor market experiences.

Figure 2-1: Timeline of Events Related to the Headscarf Ban



As mentioned in the first section, a headscarf ban is also a relevant topic for other countries. For instance, Tunisia also has banned the headscarf in public schools and universities or government buildings since 1981 (Dunbar, 2009), whereas Syria banned full face veils in universities from July 2010 (Chick, 2010). This issue is also a hotly debated topic in Europe. At the EU level, the European Court of Human Rights has ruled that headscarves might legitimately be restricted in EU countries (Vakulenko, 2007). On the other hand; Franco Frattini, who was the European Commissioner in 2006, has said that he was not in favor of banning full-face veils[7]. At the government level, France is the first country in Europe to implement a headscarf ban since September 2004 with law no 2004-228. It bans wearing all conspicuous religious symbols in French public primary and secondary schools.

In Belgium, some municipalities apply a ban only to full-face veils (Mardell, 2006). Alain Destexhe is a Belgian senator who proposed a bill that would ban headscarves from all state schools. There were two incidents about full-face veils in the UK. For one of the cases, the House of Lords stressed that this judgment cannot be generalized to address whether Islamic dress is allowed or not in UK. Therefore, we can say that in general, headscarves are allowed in the UK. In Germany, female Muslim teachers wearing headscarves have become an issue (Human Rights Watch, 2009). Norway

[7] http://www.refdag.nl/nieuws/binnenland/brussel_tegen_boerkaverbod_1_192669

has interpreted the headscarf ban as violating its "Gender Equality Act" (Skjeie, 2006). In the Netherlands, although in general, government allows its employees to wear a headscarf without much ado, there still have been some controversies about the acceptability of the headscarf in both public institutions and private enterprises (Saharso, 2007). Some of these cases were brought before the Commission of Equal Treatment. In December 2005, in nearly all cases where it has been consulted, the Commission has ruled that wearing headscarf cannot be banned because it violates the Dutch anti-discrimination law. The extreme-right parliamentarian Geert Wilders suggested the Minister of Justice to implement a ban of wearing headscarves for all public officers, yet it was rejected (Saharso, 2007). Currently, the headscarf is not banned in the Netherlands.

Both France and Turkey banned use of headscarves as discussed earlier. However, the differences between Turkey and France in terms of the potential effect of the ban are considerable. Firstly, in France, only public schools are affected by this provision. However, in Turkey, all types of schools, including private schools, are in the coverage of the ban. Moreover, in France, the ban is only applied in primary and secondary schools, which might not really affect educational attainment of Muslim women, because according to religious rules, females are supposed to wear it when they enter into adolescence. However, in Turkey, the ban is also applied in tertiary education institutions, which means that some conservative female adults may not continue their education if they do not want to uncover. Furthermore, in France, on some occasions the costs of private schooling of students who would not accept the ban on religious symbols were thus paid for by the state rather than those families. In addition, the French government operates a distance learning agency, the CNED, which is another solution for families impacted by the rules of public schools. Distance education is also an option in Turkey, but that is not trouble-free for women with headscarves either[8].

2.3 The Secondary and Tertiary Education System of Turkey

In Turkey, secondary education consists of three-year general high schools (after 2009, four years) and three or four-year vocational high schools. In secondary schools where English is the language of education, there is one additional year of language preparation. General high schools offer a

[8] There were two incidents that girls wearing headscarf were taken out during Distance Education Exam in 2009 according AK-DER headscarf ban chronology records.

curriculum preparing students for university education, whereas vocational high schools offer technical education preparing students for vocational higher education within the tertiary system. Tertiary education is composed of two or four-year vocational higher education programs and four-year (six years for medicine) programs that grant undergraduate degrees.

There exists an excess demand for tertiary education. Therefore, high school students in their final year take part in centrally administered competitive national examination to enter a university. As of 2011, 1,759,998 students have applied to University Entrance Examination, and 789,167 have been granted admission, which corresponds to 44.8%. (MSPC, 2011). Since entrance to tertiary education institutions is highly competitive, as of 2011, 4,170 private tutorial centers operate all over the country and prepare high school students explicitly for university entrance examination (Ministry of Education, 2011). Usually, high school students attend private tutorials after school or at weekends for sixteen hours a week on average. Another statistic that tells the importance of private tutorial centers is that as of 2002, 4.47% of all educational expenditure, including public educational expenditure, goes to private tutorial centers (TURKSTAT, 2002). Moreover, 10.07% of all educational expenditure by households is on private tutorial centers for university placement exam preparation in 2002 (TURKSTAT, 2002). As of 2011, one year registration to private tutorial centers costs between 1,500 TL [€ 750] to 3,500 TL [€1,750].

Another aspect of the cost of getting a university degree is tuition fees to be paid. Tuition fee levels in public institutions are centrally set, and all universities charge the same amount of tuition fee for the same programs. Private universities are free to determine their tuition fees. Student and parental contribution to the tuition fees of tertiary education in public institutions in Turkey ranges from 950 TL [€475] to 4,100 TL [€ 2,050].

The second component of the cost of tertiary education is living expenses. A sizeable portion of the students are also eligible for living in the subsidized public dormitories, where they only pay about one-third of the total cost (Eurydice, 2008). For the 2005-2006 academic year, The International Comparative Higher Education Finance and Accessibility Project (ICHEFAP, 2009) estimated tertiary education expenses born by parents and students including living and school related expenses as 2,673 TL [€ 1,337] for public universities and 9,860 TL [€4,930] for private universities on average. That is,

entrance to a university and obtaining a degree requires a significant amount of time and monetary investment on the level of student as well as by the parents.

2.4 The Impact on Female Tertiary Educational Attainment

In this section, we explore the effect of the ban by comparing tertiary educational indicators with Turkey's neighbors and also comparing female tertiary educational attainment figures with that of males in Turkey and lastly we look the effect of the ban on women wearing headscarves.

2.4.1 Aggregate Country Level Analysis

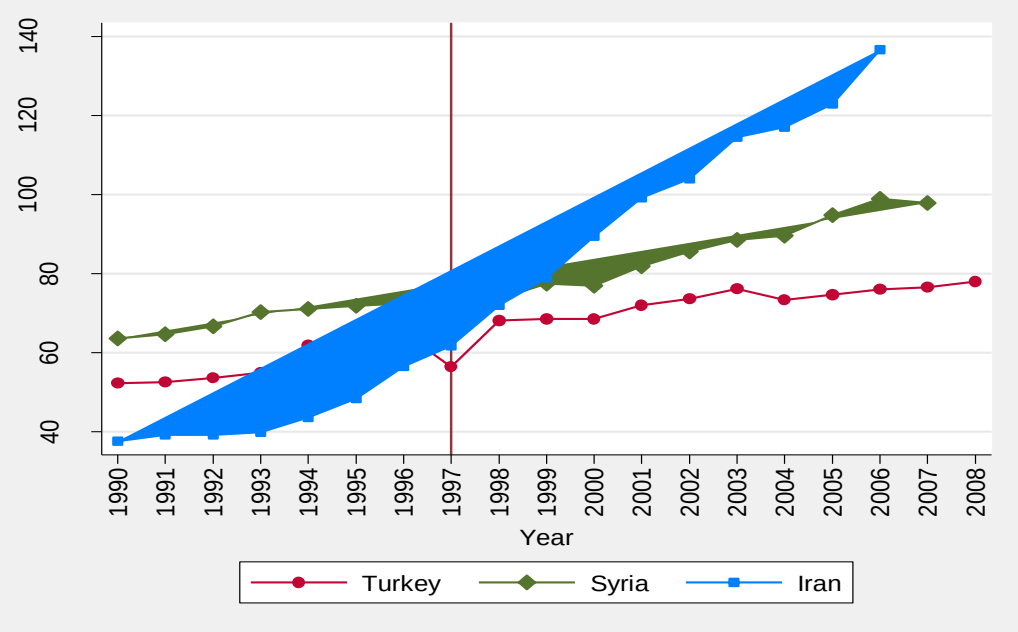
To get an overall impression of the impact of the ban, we need an appropriate control group. We use Turkey's neighbors Iran and Syria[9] as a control group because of geographic proximity, similar GDP per capita figures as of 1990, and the majority Muslim population. In terms of geographic proximity, one can also think of Greece and Bulgaria, however, GDP per capita in Greece was almost four times higher than Turkey as of 1990. Although Bulgaria is also a neighbor and had comparable GDP per capita figures as of 1990, it is a predominantly Orthodox-Christian country (CIA, 2012). Iran's GDP per capita is very close to that of Turkey's. Syrian per capita GDP is much lower, but one similar aspect is that both Turkey and Syria are predominantly Sunni-Islam countries while Iran is Shia-Muslim country. (CIA, 2012). We focus on the female to male ratio in total number of tertiary education students from Iran and Syria for 1990-2008 period. The female to male ratio in total number of tertiary education students for all countries is obtained from World Bank's database[10] to avoid incomparability[11] whenever possible. Although the data from different countries might not be plagued with different data definitions and sample selection procedures, comparing countries is essential for our analysis and using the available data is our best option.

[9] Egypt could not be included because there were only two observations for Egypt in 1990 and 1991.

[10] World Bank refers to United Nations Educational, Scientific, and Cultural Organization (UNESCO) Institute for Statistics as source of data. UNESCO refers to school register, school survey or census for data on enrolment by level of education; population census or estimates for school-age population for calculation of the statistics. Again, these figures are also officially submitted data by national authorities.

[11] UNESCO claims that it provides well-defined standards to ensure data comparability in national and international education statistics (UNESCO, 2012) .

Figure 2-2: Female to Male Ratio in Tertiary Education Students



Source: World Development Indicators, World Bank

Figure 2-2 presents female to male ratio in number of tertiary education students in Turkey, Syria and Iran before and after the headscarf ban enactment. According to Figure 2, pre-1997 trends in female to male ratio in tertiary education students in Iran and Syria are parallel to the pre-1997 trends in Turkey. Table C-1 in Appendix C provides descriptive statistics of the ratio of females to males in tertiary education students for the countries we analyze. According to Table C-1, Turkey fares better before 1997. However, it falls back after 1997 period.

We run fixed effects models to see the significance of the descriptive statistics. The identifying assumption for the results provided in Table 2-1 is that Turkey, Iran and Syria would have followed a parallel path in female/male students in tertiary education in the absence of the headscarf ban. In other words, the growth in ratio of females to males in tertiary education would have been the same in these countries, had no headscarf ban been introduced.

According to model (1), the headscarf ban resulted in a 27%[12] drop in the female to male ratio in tertiary education students in Turkey compared to Iran and Syria. When country specific time trends[13] are added in model (2), the effect of the ban has become statistically insignificant. In model (3), we estimate the model with a common trend for Syria and Turkey, since there is no marked difference

[12] 1-exp(-0.32)=0.27
[13] Time-trend of Iran is statistically different from trend of Turkey and Syria.

between Syrian and Turkish trend in model (2). In model (4), the lagged effect of the ban is estimated. Although the ban is estimated to lead to a drop of 2% in tertiary education students' female to male ratio, the effect is not statistically significant.

Table 2-1: Fixed Effects Estimates of log Female/Male Ratio

	Total Number of Students			
	(1)	(2)	(3)	(4)
Ban	0.54***	0.05	0.04	
	(8.13)	(1.32)	(1.28)	
Ban*Treatment	-0.32***	-0.06	-0.04	
	(-2.90)	(-1.04)	(-1.16)	
Lagged Ban				0.09***
				(2.77)
Lagged Ban*Treatment				-0.02
				(-0.61)
Trend-Iran		0.09***	0.09***	0.09***
		(20.57)	(23.11)	(22.05)
Trend-Turkey		0.02***		
		(5.99)		
Trend-Syria		0.02***		
		(5.59)		
Trend-(Turkey&Syria)			0.02***	0.02***
			(8.25)	(6.42)
Country Dummies	+	+	+	+
N	52	52	52	48
R-sqr	0.61	0.97	0.97	0.97

2.4.2 Analysis of Aggregate National Data

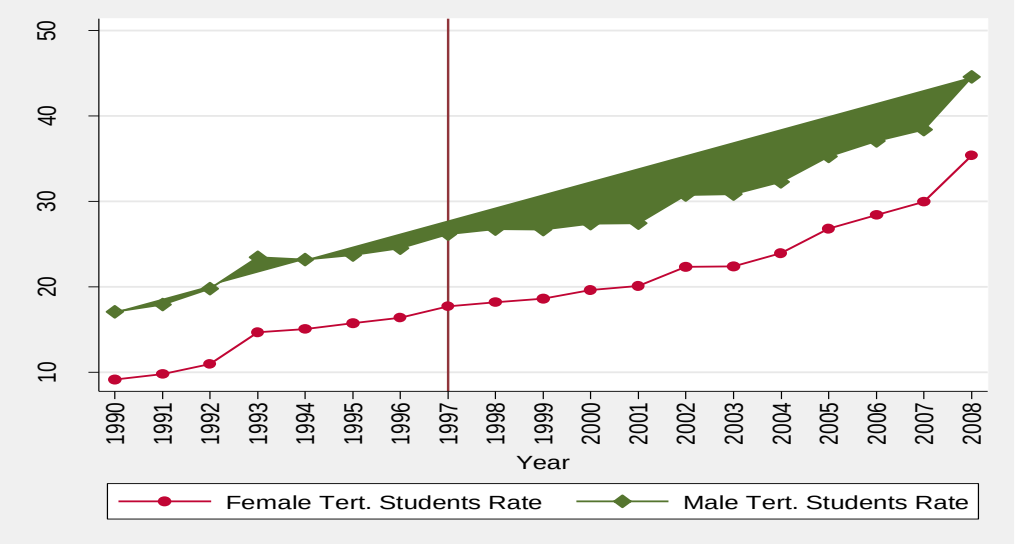
One can also study the impact of the ban on female educational attainment using males as a control group. Figure 2-3 shows the gross tertiary schooling rate[14] for males and females in Turkey. Figure 2-4 shows new admissions and graduation rates[15], which represent the flow in and out of tertiary education. According to Figure 2-3 and 2-4, pre-1997 trends for males are comparable to that of females.

According to Figure 2-3 and 2-4, even after 1997 period, the female and male gross schooling rates as well as female and male new admissions rates continued to follow a similar path, which suggests that the ban did not impact female new admissions considerably. However, the female graduate rate has leveled off for 4 years, starting with 1996/97 academic year.

[14] Gross tertiary schooling rate is defined as total number of students in a tertiary education as a percentage of 20-24 year old population.

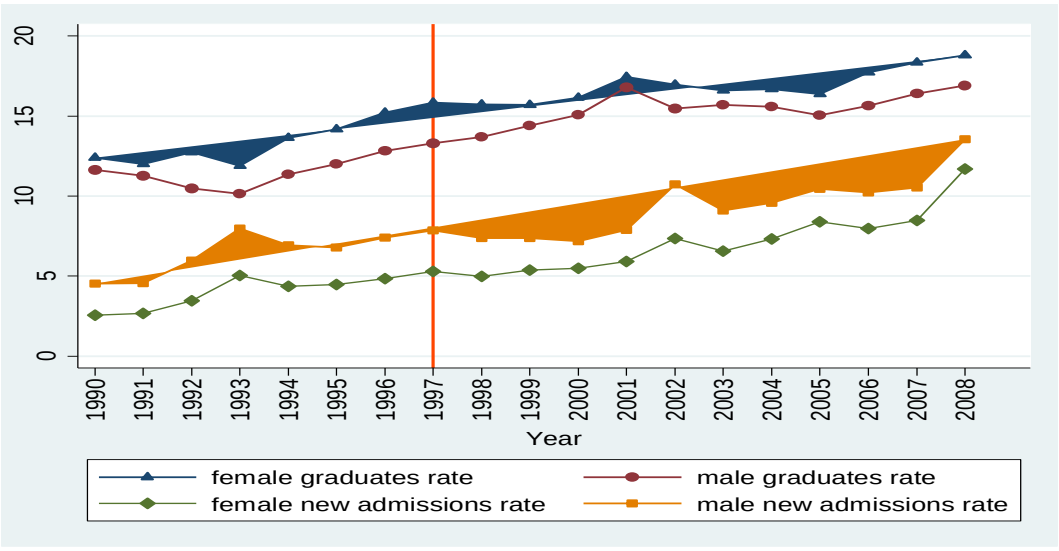
[15] New admissions rate is calculated as dividing total new admissions by 20-24 year old population figures and graduate rate is calculated as dividing total number of graduates by total number of tertiary students respectively.

Figure 2-3: Tertiary Education Indicators in Turkey



Source: National Education Statistics, TURKSTAT

Figure 2-4: New Admissions and Graduates of Tertiary Education Institutions



Source: HEC Stastics Yearbooks and National Education Statistics, TURKSTAT

In Table 2-2, fixed effects estimates of total number of students and graduates rates[16] are displayed. According to model (1), the headscarf ban resulted in a 7% drop female tertiary education students compared to males, but the effect is not statistically significant. In model (2), gender specific time trends are added. Again, the effect of the ban is statistically insignificant. In model (3), we estimate the model with a common trend for males and females, since there is no marked difference between female and male trend in model (2). Similar models are also run for graduates rates. We did

[16] Similar analysis is conducted for new admissions rate, but for the sake of saving space, they are not displayed in the Table 2-2 since the results are not much different from total number of students regression results.

not find any statistically significant effect of the ban on females compared to males in Table 2-2. Since there is excess demand for tertiary education, females wearing headscarves could easily be replaced by those who are not wearing headscarves. In short, we did not detect any effect of the ban on aggregate tertiary education indicators using neighboring countries as control group and males as control group.

Table 2-2: Fixed Effects Estimates of Tertiary Education Indicators

	Total Number of Students			**Graduates Rates**		
	(1)	(2)	(3)	(4)	(5)	(6)
Ban	10.59***	-3.54**	-3.31***	3.96***	1.71***	1.50***
	(4.55)	(-2.54)	(-2.99)	(7.72)	(2.74)	(3.02)
Ban*Treatment	-0.07	0.39	-0.07	-0.24	-0.65	-0.24
	(-0.02)	(0.20)	(-0.07)	(-0.33)	(-0.74)	(-0.50)
Trend-Female		1.49***			0.24***	
		(12.14)			(4.32)	
Trend-Male		1.44***			0.28***	
		(11.74)			(5.11)	
Trend-			1.46***			0.26***
(Female & Male)			(17.12)			(6.74)
Country Dummies	+	+	+			
N	38	38	38	38	38	38
R-sqr	0.55	0.95	0.95	0.77	0.90	0.90

2.4.3 Analysis of Micro Data

Although analyzing aggregate national data before and after the ban allows for identification of the impact of the ban, we cannot study the effects on specific population groups with aggregate data. Indeed, this ban is expected to impact the outcomes for women who prefer to wear the scarf, and it should not impact women who do not wear it. We also undertake an analysis to determine whether the ban had any impact for women who prefer to use the headscarf. We only have individual level data for headscarf use status after the ban. Table 2-3 provides the summary of the surveys that we use for individual level analysis. More information on the micro-data is provided in Appendix A.

Table 2-3: Summary of the Surveys

Company	Year	Type of Survey	Available Dependent Variables	Sample Size	% wearing headscarf
Konda	2007	Cross-Section	Educational Attainment, Employment	2,639	67
A&G	2003	Cross-Section	Educational Attainment, Employment	927	66.24
A&G	2007	Cross-Section	Educational Attainment, Employment	1,316	57.45
NFHS	2003	Cross-Section	Educational Attainment, Employment, Fertility	8,075	75.4
NFHS	2008	Cross-Section	Educational Attainment, Employment, Fertility	7,405	75.6

The first data that we use is from Konda Research Company's survey in 2007. The second dataset is from A&G research company's two field surveys conducted in 2003 & 2007. The third dataset is National Family and Health Surveys conducted in 2003 and 2008. The descriptive statistics are provided in Table C-2, Table C-3 and Table C-4 respectively. Appendix A provides details of the wording of wearing headscarf questions. Descriptive information provided by all the surveys indicates that there is a large educational gap between women wearing headscarf and not wearing headscarf.

To see a clearer picture of the link between headscarf use and educational attainment, potentially confounding factors need to be controlled. We start with the Konda 2007 survey. Our control variables are age in categories[17], marital status, household size, household income, current region of residence, urban/rural status, and region of birth in the regression model together with headscarf dummy variable. The dependent variable is having a university or higher education degree. Marginal effects calculated from probit models are presented in Table 2-4.

Table 2-4: Probit Estimates of Tertiary Degree Holding (Marginal Effects)– Konda 2007

	(1)	(2)	(3)	(4)
Headscarf	-0.083***	-0.073***	-0.077***	-0.081***
	(-7.60)	(-6.28)	(-6.00)	(-5.67)
Baseline Controls	+	+	+	+
Self-reported Religiosity	-	+	-	-
Religious practices	-	-	+	+
Women should be able to work	-	-	-	+
N	2,498	2,467	2,459	2,247
Pseudo R-sqr	0.35	0.35	0.36	0.35

From model (1), women wearing a headscarf are 8.3% less likely to hold a tertiary education degree. The results in model (1) might be just because there is less demand for education among religious people. In order to control for that, the individual's own reported degree of religiosity is included in model (2). Moreover, individual frequency of praying the daily five prayers, fasting, reading the Quran and making voluntary prayer are included in model (3). The coefficient might still be attributed to value structure differential. In order to reduce bias in the estimate, one can incorporate ideas on women's paid work. The answers to the question of whether women should work in order to

[17] Age categories in the survey are 18-28 years old, 29-43 years old, 44 or more years old. Because of these wide categories, we could not exploit the fact that some age groups are exposed to the ban and some are not in the regression.

contribute family budget[18] are included in model (4). The figures for models (2), (3) and (4) are also very similar. According to model (4), women wearing headscarf are 8.1% less likely to hold a university or higher degree. However, it is hard to attribute this solely to headscarf ban. It could also be that women who do not want to get higher education are more likely to use a headscarf. That is, women who use the headscarf are aware of the consequences of it in terms of their educational attainment. So, the choice to wear the scarf is endogenous. Therefore, one cannot interpret the results as the effect of the ban. Yet, the results still suggests that even after controlling for engaging in religious practices, women wearing headscarves have significantly lower educational attainment.

A similar analysis is conducted using a combined version of A&G research's surveys in 2003 and 2007. For that purpose, having a tertiary education degree or being a tertiary education student is regressed on age, marital status, current household income, current region of residence, urbanity, year dummy, whether the respondent reads the daily newspaper, age at most 17 in 1997 dummy variable, headscarf dummy variable and headscarf dummy variable interacted with age at most 17 in 1997 dummy variable. The marginal effects calculated from probit models are reported in Table 2-5.

Age at most 17 in 1997 is used as a cut-off point because these women are exposed to the headscarf ban fully. The sample is restricted to women who at most 17 years old in 1990, because 1990 marks the beginning of the relatively free period for women with headscarves. Moreover, the headscarf ban would not really matter for the education decisions of older generations. Model (2) differentiates between the potential impacts of the ban on different age-group of women. 18-21 year old women in 1997 were more likely to be in higher education institutions. This age-group would have been impacted by the ban whereas women older than 21 in 1997 could have graduated already, therefore might not be affected by the ban. We call 18-21 year old women a transition group. In model (3), we restrict the fully exposed ones to women aged 16 at most in 1997. Since entering tertiary education requires a long-preparation period, women aged 17 have possibly already invested in tertiary education preparation for a year, and thus could be grouped with women of ages 18-21 in 1997.

[18] The responses range from, "I agree", "it depends", "I do not agree", "I do not know". Therefore, 4 dummies are added for each category.

Table 2-5: Probit Estimates of Tertiary Degree Holding (Marginal Effects) -A&G

	(1)	(2)	(3)	(4)	(5)	(6)
Headscarf	-0.129***	-0.208***	-0.211***	-0.106	-0.200	-0.213
	(-3.52)	(-3.44)	(-3.58)	(-1.15)	(-1.36)	(-1.51)
Headscarf*age at most 17 in 1997	-0.027	0.049		-0.132	-0.035	
	(-0.55)	(0.70)		(-1.10)	(-0.21)	
Headscarf*age 18-21 in 1997		0.132*			0.153	
		(1.70)			(0.84)	
Headscarf*age at most 16 in 1997			0.018			-0.091
			(0.25)			(-0.54)
Headscarf*age 17-21 in 1997			0.167**			0.237
			(2.35)			(1.45)
Baseline controls	+	+	+	+	+	+
N	892	892	892	401	401	401
Pseudo R-sqr	0.30	0.30	0.33	0.15	0.16	0.20

In model (4), the sample is restricted only to respondents holding a high-school degree, tertiary degree or in the process of obtaining the tertiary degree, to check whether the headscarf ban influences only the transition from high school to tertiary education or not. Model (5) checks for transition to higher education for different age-groups as in model (2). Model (6) checks for transition to higher education, but we use the classification of model (3). The interaction term of headscarf with age categories is the variable of interest for observing the effect of the headscarf ban.

In all 6 models, the interaction term is either insignificant or positive. We did not find any evidence that the headscarf ban negatively impacted the educational attainment of women wearing headscarves. A word of caution is in order here. Firstly, the sample size is small, among them wearing headscarves fewer. Secondly, in this dataset, we do not know whether the women have graduated before the ban or not. We only have the age of the women. Not all children begin school in the year predicted by school entry policies. The parents of a child born before the school entry date may hold their child back for a year and the parents of a child born after school entry date may petition for their child to start school a year before, which is typically allowed, or they may send their child to private school, where school entry policies are less strict. As a result, compulsory age of education is not strictly enforced. There could be many same aged people going to different classes. Moreover, some schools have preparatory classes, so using age to predict when the respondent might have graduated from university is therefore a crude measure.

The third dataset used for this research is NFHS-3 combined with NFHS-4. Mother's and father's education level, region lived, type of place of residence, age, mother tongue, wealth index, household assets such as car, motorbike, TV, refrigerator, telephone as a proxy for wealth, survey year, born after 1980, headscarf use and interaction of born after 1980 and headscarf are used as baseline controls. "Born after 1980" dummy variable corresponds to "age at most 17 in 1997" dummy variable in A&G's survey. Born after 1980 is used as a cut-off point, because women born after 1980 are exposed to the headscarf ban fully. Women who are currently younger than 17 years old do not have a chance to have higher education, thus they are discarded from analysis. Moreover, educational decisions of older generations might not be comparable with younger people. Therefore, women born before 1973 are also discarded from analysis. This is the same as restricting the sample only to women at most 17 years old in 1990 in A&G's survey.

Women born before 1976, if they did not repeat any grade, could get a degree without being subject to the headscarf ban, whereas women born between 1979 and 1976 were possibly at school when the ban was enacted. In model (2), we differentiate between different possible effects on different age groups. Similarly, we include women born in 1980 into the transition group in model (3), since they might have already incurred the costs of preparing for university entrance examination.

In model (4), the sample size is restricted to respondents having a high-school degree versus tertiary education degree, to check whether the effect is more on transition to higher education or not. Model (5) checks on the effect of the ban on the transition to higher education, while differentiating the effects on different age-groups. Similar to model (3), we include women born in 1980 into the transition group and also check on the effects of ban on transition to higher education.

The marginal effects calculated from probit models are shown in Table 2-6. From model (1), women wearing a headscarf are 5.3% less likely to have a tertiary education. Model (2) also shows similar results. According to model (4), women wearing a headscarf are 12% less likely to have a tertiary education compared to having a high school degree. However, the interaction term is insignificant in all of the models; therefore, we could not find any evidence for significant change in the educational attainment of women wearing headscarves after the imposition of the ban.

Table 2-6: Probit Estimates of Tertiary Degree Holding (Marginal Effects) - NFHS-2003&2008

	(1)	(2)	(3)	(4)	(5)	(6)
Headscarf	-0.053***	-0.052***	-0.052***	-0.116***	-0.125**	-0.125**
	(-7.05)	(-4.64)	(-4.63)	(-3.64)	(-2.49)	(-2.50)
Headscarf*Born after 1980	-0.001	-0.001		0.025	0.033	
	(-0.05)	(-0.04)		(0.51)	(0.53)	
Headscarf*Born between 1976&79		-0.000			0.016	
		(-0.02)			(0.25)	
Headscarf*Born after 1981			0.006			0.049
			(0.41)			(0.75)
Headscarf*Born between 1976&80			-0.005			0.008
			(-0.33)			(0.13)
Baseline Controls	+	+	+	+	+	+
N	6,593	6,593	6,593	1,428	1,428	1,428
Pseudo R-sqr	0.37	0.38	0.38	0.13	0.13	0.13

2.5 The Impact on Labor Force Participation of Women

In this section, the impact of the ban on female employment is examined firstly in aggregate terms using neighboring countries as control group and secondly males as control group. Thirdly, an analysis is conducted on women wearing headscarves using micro data.

2.5.1 Aggregate Country Level Analysis

In this section, we analyze the impact of the ban for average FLFP and using Turkey as treated unit and its neighbors Iran, Syria and Egypt as counterfactuals. The FLFP rate for all countries is obtained from World Bank's database to avoid incomparability[19].

Table C-5 provides descriptive statistics for these indicators before and after the headscarf ban enactment. The average FLFP rate was higher in Turkey than that of the control group before 1997 which was for the same period. After 1997 period, Turkey's FLFP rate fell down whereas the control group's FLFP has increased slightly. We analyze the statistical significance of the information provided in Table C-5 using difference-in difference methodology. Table 2-7 provides the estimation results.

[19] Since these countries are very different in many aspects, comparability of statistics is of a considerable concern. Therefore, we use World Bank's database which obtained FLFP rate figures from International Labor Organization's Key Indicators of the Labor Market database. ILO claims that the indicators are to a large extent comparable across countries since they use standardized indicators. Another thing that suggests comparability across countries is that ILO indicators rely heavily on the official submission of data by national authorities (International Labor Organization, 2013).

Table 2-7: Fixed Effects Estimates of log Female Labor Force Participation

	(1)	(2)	(3)	(4)
Ban	0.06	0.02	0.02	
	(1.42)	(0.48)	(0.58)	
Ban*treatment	-0.25***	-0.01	-0.01	
	(-2.85)	(-0.06)	(-0.25)	
Lagged Ban				0.02
				(0.57)
Lagged Ban				0.03
*Treatment				(0.51)
Trend-Turkey		-0.02***		
		(-3.16)		
Trend-Syria		-0.02***		
		(-4.12)		
Trend-Iran		0.04***	0.04***	0.04***
		(8.22)	(8.54)	(8.80)
Trend-Egypt		-0.01	-0.01	-0.00
		(-1.49)	(-1.59)	(-0.84)
Trend- Turkey&			-0.02***	-0.02***
Syria			(-5.23)	(-6.08)
Country Dummies	+	+	+	+
N	76	76	76	72
R-sqr	0.10	0.74	0.74	0.77

Model (1) indicates that FLFP has dropped by 22%[20] in Turkey compared to the control countries, and the effect is statistically significant. In model (2), when country specific time trends are controlled for, the effect is no longer statistically significant. Time-trends for Iran and Egypt are markedly different from that of Turkey and Syria. However, there is no statistically significant difference between time-trends of Turkey and Syria, therefore in model (3), we estimate a common trend for Turkey and Syria. Model (4) estimates the lagged effect of the ban. In both model (3) and model (4), the effect of the ban is statistically insignificant.

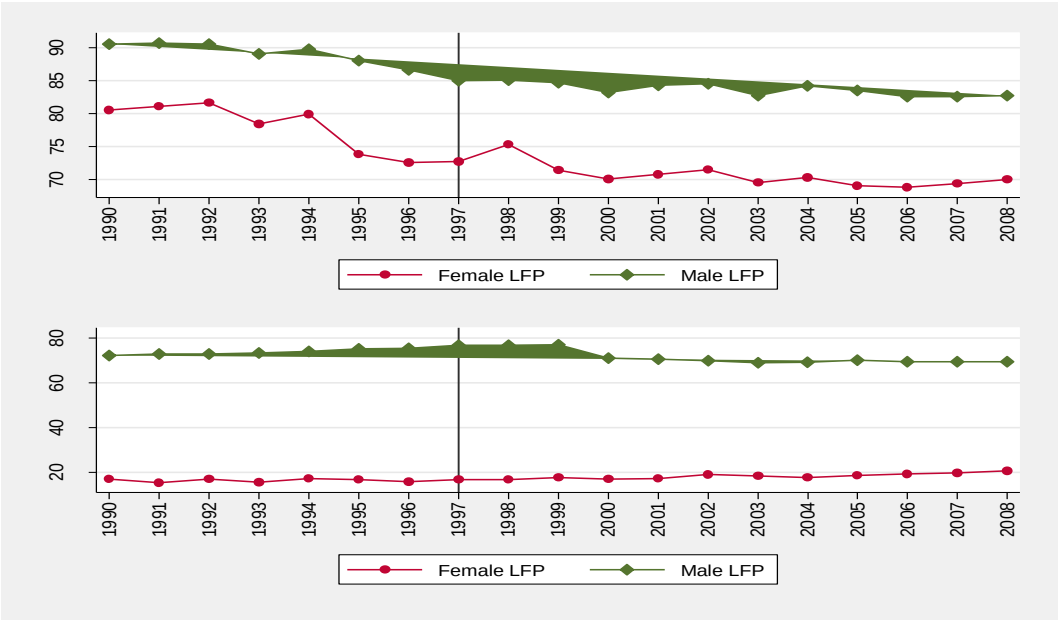
2.5.2 Analysis of Aggregate National Data

In this section, we study the impact of the ban on female labor market outcomes using males as a control group. The headscarf ban and its implications for the labor market are more of an issue for higher educated and in urban areas. Therefore, we firstly focus on urban labor market figures because the overall labor market figures are also affected by the rural labor market. Rural labor market is mainly driven by the agricultural sector, where it is less likely to observe any effect of the headscarf

[20] 1-exp(-0.25)=0.22

ban because agricultural holdings are not institutionalized, and generally women work as unpaid family workers. For having a general feeling of the context, we start with urban LFP and LFP of higher educated for each sex depicted in Figure 2-5.

Figure 2-5: Labor Force Participation Rates by Gender in Turkey

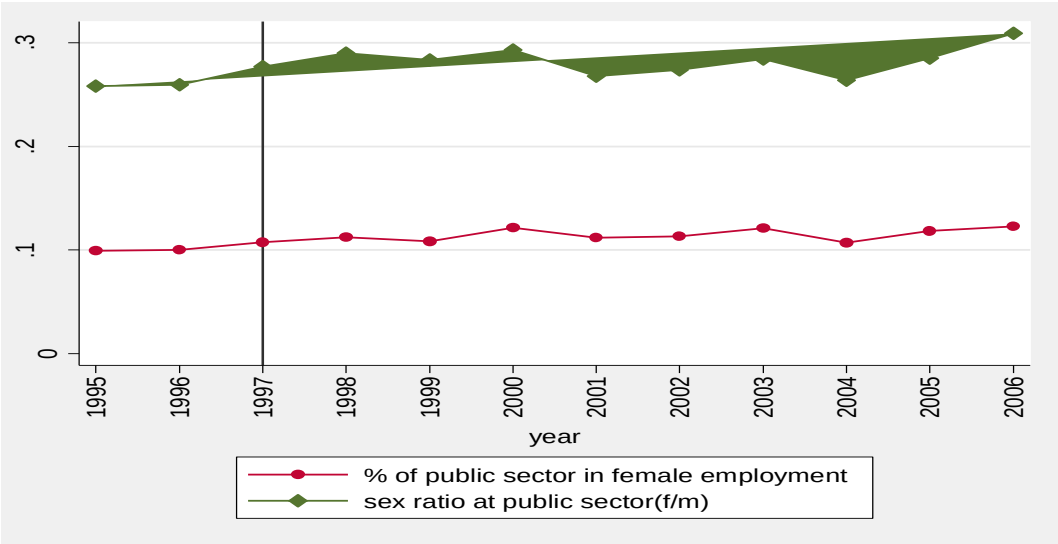


Source: Gender Indicators, TURKSTAT

Despite the improvement in legislature aimed at removing gender discrimination in employment as of 2003, UFLFP has not changed much over the years. Moreover, we do not see a large change after 1997 for female LFP compared to males. It can also be seen from Figure 2-5 that the labor force participation (LFP) of higher educated women has been declining over the years. The LFP of higher educated men also fell slightly. Yet, when compared with the figures for males, no abrupt change is visible in LFP of higher educated women.

Since the headscarf ban directly impedes public sector employment opportunities for a large number of women, we shall examine public sector employment of women. Figure 2-6 displays statistics regarding female public sector employment. The series is available only from 1995 to 2006 through TURKSTAT Household Labor Force Surveys. The ratio of female to male employees in the public sector was 25% in 1995; this ratio fell down in 1999 and 2001 slightly. Overall, the sex ratio did not change much over time. Similarly, the gender ratio in the public sector stayed almost the same over time. These two indicators do not show an abrupt change in female employment in the public sector after 1997.

Figure 2-6: Public Sector Employment of Women



Source: Household Labor Force Surveys, TURKSTAT

In Table 2-8, we provide the estimates of the effect of the ban on urban LFP and LFP of higher educated females using males as control group.

Table 2-8: Fixed Effects Estimates of Labor Market Outcomes

	Urban LFP Rate		**LFP of Higher Educated**		
	(1)	(2)	(3)	(4)	(5)
Ban	-0.02**	0.03*	-0.06***	-0.03**	-0.02
	(-2.18)	(1.97)	(-5.73)	(-2.13)	(-1.41)
Ban*Treatment	0.04***	-0.03*	-0.02	0.01	-0.02*
	(2.86)	(-1.70)	(-1.46)	(0.29)	(-1.92)
Trend-Female		-0.01***		-0.00**	
		(-4.15)		(-2.49)	
Trend-Male		0.00**		-0.01***	
		(2.04)		(-4.79)	
Trend-					-0.00***
(Female & Male)					(-5.02)
Country Dummies	+	+	+	+	
N	38	38	38	38	38
R-sqr	0.19	0.52	0.55	0.95	0.95

According to model (1), the headscarf ban resulted in a 4% rise in urban labor force participation rate for females compared to males. When gender specific time trends[21] are added in model (2), we observe a 3% drop in urban LFP rate for females and the effect is statistically significant at 10% level. Similar models are also run for LFP of higher educated women. Model (1) suggests a 2% drop in LFP of higher educated women after the ban compared to higher educated males LFP. But, the effect is not statistically significant. When gender specific time trends are introducted in model (2), we also

[21] The gender specific time trends are markedly different from each other, therefore a model with a common trend is not estimated.

observe an insignificant effect of the ban. However, since there is no marked difference between females and males time trends, we estimate model (3) with a common trend for males and females. According to model (3), the ban resulted in 2% drop in LFP of higher educated women and the effect is statistically significant at 10% level.

2.5.3 Analysis of Micro Data

Table C-6 provides descriptive information from Konda's 2007 survey on labor market outcomes for women by the use of headscarf. Table C-7 and Table C-8 also provide similar descriptives using A&G's surveys conducted in 2003 & 2007 and NFHS-2003 & NFHS-2008 datasets respectively. From all descriptive statistics, there are sharp and statistically significant differences in all job types and overall employment status between women wearing headscarves and not wearing the scarves. Women wearing headscarves are less likely to be employed. When one looks into the categories of employment, women wearing headscarves are less likely to be employed in the public sector and the private sector, and are also less likely to be self-employed.

It is not surprising that there is a difference between women wearing the headscarf and the ones that do not in public sector jobs. However, there are also very sharp differences in the private sector as well. Cindoglu (2010) argues that the ban has a spill-over effect. Private companies do not prefer to hire women with headscarves, because they cannot do their job whenever there is a contact with public offices. Those women have to be invisible in offices. Due to the nature of white-collar jobs, a journalist, an engineer, or a banker would have to deal with many different institutions, some of which impose a ban on the headscarf. Encounters with such institutions may result in poor performances, which may make a woman with headscarf a liability for the company. Cindoglu (2010) suggests that even if there are no discriminatory motivations or intentions from the company, the existence of the headscarf ban creates a hostile environment for professional women who want to wear the scarf. According to Konda's 2007 survey results, there is a wide gap between the LFP rate of women with tertiary education by their use of headscarf. Among women with tertiary education, 35% of women wearing headscarves are employed, whereas 68% of women without the scarf are employed.

We also conducted a regression analysis. We first look into Konda's 2007 survey. Our baseline control variables are age, marital status, household size, household income, current region of residence, urban/rural status, region of birth together with headscarf dummy in model (1). In model (2), we include education status. One can argue that there might be less demand for employment among religious women. In order to control for that, individual's own reported degree of religiosity is included in model (3). Instead of religiosity, we included individual frequency of praying, fasting, reading the Quran and making voluntary prayer as additional regressors in model (4). We also tried to incorporate a value structure by including the responses to questions on women's work. In Konda 2007 survey, the respondents are asked whether they agree with the following statement; 'women should work in order to contribute family budget'. The responses to this question are also included in the model (5)[22]. Marginal effects calculated from probit models are reported in Table 2-9. Although according to model (1), women wearing headscarf are 12.3% less likely to be employed; the effect of headscarf goes down in models (2), (3) and (4). Model (5) suggests that women wearing a headscarf are 4.1% less likely to be employed after accounting for extensive set of controls.

Table 2-9: Probit Estimates of Employment Status (Marginal Effects)– Konda 2007

	(1)	(2)	(3)	(4)	(5)
Headscarf	-0.123***	-0.079***	-0.075***	-0.045**	-0.041**
	(-7.58)	(-4.70)	(-4.16)	(-2.29)	(-2.06)
Baseline Controls	+	+	+	+	+
Education Level	-	+	+	+	+
Self-reported Religiosity	-	-	+	-	-
Religious practices	-	-	-	+	+
Women should be able to work	-	-	-	-	+
N	2,504	2,496	2,465	2,457	2,455
Pseudo R-sqr	0.14	0.17	0.17	0.18	0.18

Using A&G's surveys conducted in 2003 and 2007, employment status is regressed on headscarf use, age, marital status, current household income, current region of residence, urbanity, survey year, whether the respondent reads a newspaper in model (1). Model (1) is the baseline model. In the model (2), we also included education status. Prayer frequency was only asked in the 2003 survey. In model (3), the baseline regressors plus praying frequency are used in explaining employment status. In model (4), education status is added to the model (3). The regression results are summarized in Table

[22] The responses range from, "I agree", "it depends", "I do not agree", "I do not know". Therefore, 4 dummies are added for each category.

2-10. According to model (1), women wearing a headscarf are 12.9% less likely to be working, but when education and prayer frequency is incorporated in model (4), women wearing a headscarf are 4.9% less likely to be working.

Table 2-10: Probit Estimates of Employment Status (Marginal Effects) –A&G 2003& 2007

	(1)	(2)	(3)	(4)
Headscarf	-0.129***	-0.084***	-0.078***	-0.049*
	(-8.05)	(-5.07)	(-2.92)	(-1.84)
Baseline Controls	+	+	+	+
Education	-	+	-	+
Praying Frequency	-	-	+	+
N	2,095	2,095	857	857
Pseudo R-sqr	0.23	0.27	0.23	0.26

The third regression analysis is conducted using NFHS 2003 and NFHS 2008 surveys. In model (1), the regressors for explaining employment status are age, mother's and father's education level, region lived, type of place of residence (urban/rural), mother tongue, wealth index and some household assets (car, motorbike, TV, refrigerator, telephone) as a proxy for wealth, survey year. Model (1) is our baseline model. In model (2), we included education status. In model (3), prayer and fasting habits are added to model (2). Prayer and fasting habits were only asked in the NFHS-2008 survey, therefore the numbers of observations are much smaller. In model (4) the ideas on whether women should work are added to the regressors used for model (3). The regression results are provided in Table 2-11.

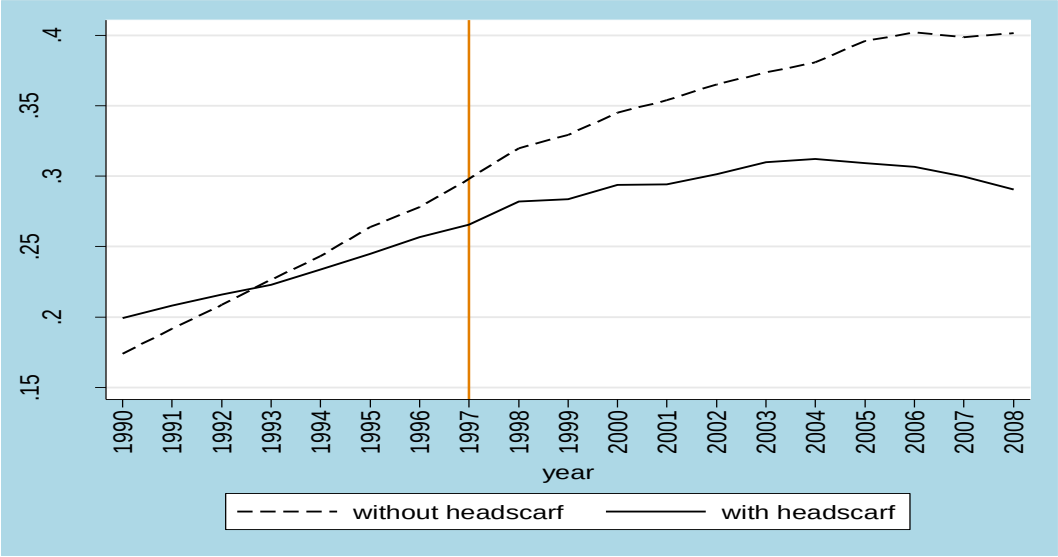
Table 2-11: Probit Estimates of Employment Status (Marginal Effects) – NFHS-2003&2008

	(1)	(2)	(3)	(4)
Headscarf	-0.104***	-0.071***	-0.032**	-0.031**
	(-11.28)	(-7.39)	(-2.08)	(-1.97)
Baseline Controls	+	+	+	+
Education	-	+	+	+
Religious Practices	-	-	+	+
Women should be able to work	-	-	-	+
N	13,880	13,880	6,710	6,650
Pseudo R-sqr	0.07	0.08	0.10	0.10

When more controls are added to the model, the coefficient of headscarf drops down from 10.4% to 3.1%. Model (4) being the most comprehensive model tells us that women wearing a headscarf are 3.1% less likely to be employed. Again, wearing a headscarf being a choice plagues the regression results with endogeneity.

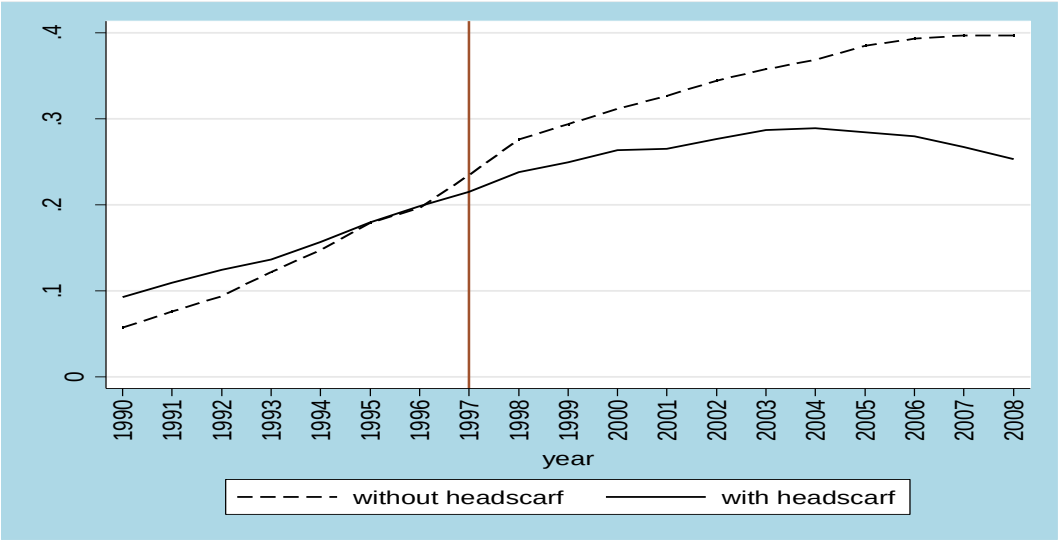
Finally, the National Family and Health survey conducted in 2008 has full employment history data. We used NFHS-2008 to construct employment figures for previous years for women with headscarves and without headscarves. The average FLFP rate for the two groups in the years 1990-2008 are shown in Figure 2-7. The headscarf ban might be more relevant for younger cohorts as we discussed for tertiary education. In Figure 2-8, we restricted the sample to only women born after or in 1973.

Figure 2-7: Female LFP over Time by Headscarf Use



Source: NFHS-2008

Figure 2-8: Female LFP over Time by Headscarf Use (younger cohorts)



Source: NFHS-2008

From Figure 2-7, we can see that both women wearing headscarves and not wearing headscarves' labor force participation has increased over time until 2004. However, the increase for women not

wearing the scarf was steeper. Labor force participation of women wearing headscarves leveled off in 2004 and then started to decline. Similarly, Figure 2-8 shows that among younger cohorts, both women with the scarf and without the scarf experienced a rise in labor force participation, although the rate of increase was higher for women not wearing a headscarf. Similarly, for younger cohort women wearing headscarves, labor force participation started to drop from 2004 onwards. However, this difference could be because of the age structure difference between women wearing headscarves and not wearing headscarves or place of residence.

The NFHS-2008 dataset has migration history, marriage history as well as the full fertility history for women. After controlling for these observable factors, we check whether there is any change in LFP of headscarved women after 1997 period. In model (1), employed dummy variable is regressed on age, marital status, current region of residence, urbanity, number of children under 1 years old, number of children under 5 years old, number of children under 16 years old, headscarf dummy variable, after 1997 dummy variable, interaction of headscarf dummy variable with after 1997 dummy variable. This is the baseline model. To capture trending, time trend is included in model (2). In model (3), education level is added to the list of regressors. For younger cohorts, we restricted the sample to women born in or after 1973. Similar models are run, initially with the baseline regressors. This corresponds to model (4) and we added time trend in model (5). In model (6), we also controlled for education status for younger cohorts.

Table 2-12: Probit Estimates of Employment Status (Marginal Effects) – NFHS-2008

	All women			Younger Cohorts		
	(1)	(2)	(3)	(4)	(5)	(6)
Headscarf	-0.029***	-0.029***	-0.002	0.026***	0.026***	0.020***
	(-6.20)	(-6.23)	(-0.34)	(4.19)	(4.21)	(3.13)
Headscarf*The ban	-0.048***	-0.048***	-0.048***	-0.059***	-0.059***	-0.059***
	(-8.76)	(-8.74)	(-8.78)	(-8.31)	(-8.29)	(-8.26)
Baseline Controls	+	+	+	+	+	+
Time Trend	-	+	+	-	+	+
Education	-	-	+	-	-	+
N	138,933	138,933	138,933	77,747	77,747	77,747
Pseudo R-sqr	0.11	0.12	0.11	0.14	0.14	0.14

Table 2-12 presents the estimated marginal effects. From model (1), women wearing a headscarf are 4.8% less likely to be employed after 1997. The figures for model (2) and (3) also show that

women wearing a headscarf are 4.8% less likely to be employed after 1997. These figures imply that after the implementation of the ban, women wearing the headscarf became less likely to be employed compared to before the ban period. As we expect, the headscarf ban is more of an issue for younger cohorts. When we look at model (6), women wearing headscarf are 5.9% less likely to be employed after the ban.

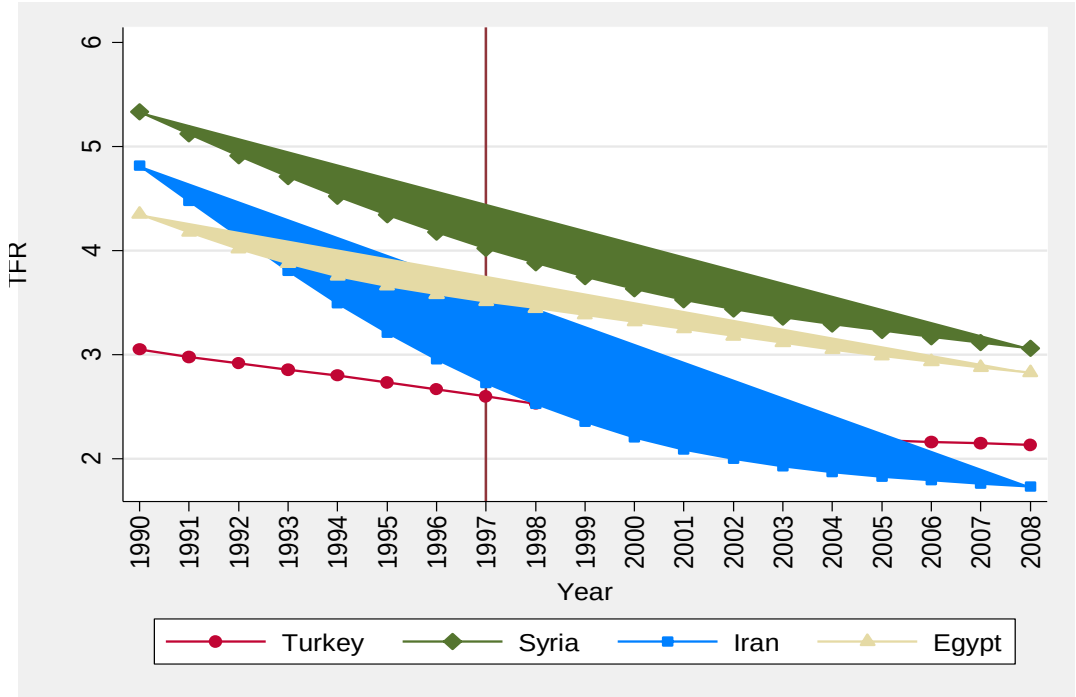
2.6 The Impact on Childbearing

Since there is a well-documented literature on the link between employment, education and childbearing, in this section we examine whether the headscarf ban affects the childbearing rate of women. For that purpose, we examine aggregate data of total fertility rate and individual level data from NFHS-4 in 2008, since it has full reproduction, employment, marriage, and migration history[23].

2.6.1 Analysis of Aggregate Data

We start our country level analysis by exploring total fertility rate for Iran, Syria, Egypt and Turkey. According to Figure 2-9, pre-1997 trends of Turkey are comparable to that of Syria and Egypt. Iran seems to have a steeper downward trend before 1997.

Figure 2-9: Total Fertility Rate by Country



Source: World Development Indicators, World Bank

[23] Although NFHS-3 conducted in 2003 also has fertility history, it does not have history of other control variables. Therefore, we restricted our analysis to NFHS-2008.

Total fertility rate figures for all countries are obtained from the same source (World Bank database)[24] to minimize potential differences in data definition and other data collection differences[25]. Table C-9 provides descriptive statistics of the total fertility rate before and after the headscarf ban enactment for these countries and reports a 0.71 higher TFR in Turkey compared to control group after the enactment of the ban. To check whether an 0.71 increase in TFR is statistically significant or not, fixed effects regression analysis is conducted. Table 2-13 provides the estimation results.

Table 2-13: Fixed Effects Estimates of Total Fertility Rate

	(1)	(2)	(3)	(4)
Ban	-1.27***	-0.32***	-0.34***	
	(-12.41)	(-3.70)	(-4.27)	
Ban*treatment	0.71***	0.20	0.27**	
	(3.48)	(1.12)	(2.50)	
Lagged Ban				-0.26***
				(-3.40)
Lagged Ban*treatment				0.20*
				(1.97)
Trend-Turkey		-0.05***		
		(-3.43)		
Trend-Syria		-0.10***	-0.10***	-0.10***
		(-10.43)	(-10.70)	(-10.34)
Trend-Iran		-0.14***	-0.14***	-0.14***
		(-14.86)	(-15.31)	(-14.24)
Trend-Egypt		-0.05***		
		(-5.64)		
Trend-Turkey&			-0.05***	-0.05***
Egypt			(-6.61)	(-6.41)
Country Dummies	+	+	+	+
N	76	76	76	72
R-sqr	0.70	0.94	0.94	0.93

According to model (1), the 0.71 higher TFR rate in Turkey is statistically significant. In model (2), when country specific time trends are controlled for, the effect is no longer statistically significant. In model (3), we estimate the model with a common trend for Turkey and Egypt, since there is no statistically marked difference between Turkey's and Egypt's trend in model (2). Model (3) indicates that the headscarf ban led to 0.27 higher TFR. In model (4), the lagged effect of the ban is estimated. Similarly, the ban had increased TFR by 0.2 children per woman.

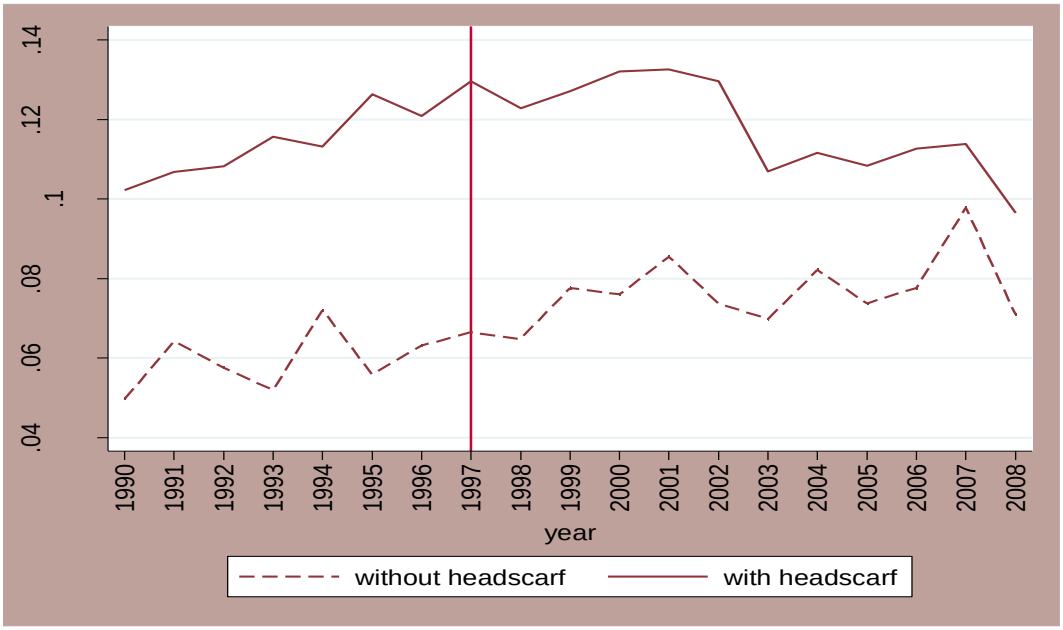
[24] World Bank Database refers United Nations (UN) Population Division, Census reports and other statistical publications from national statistical offices and U.S. Census Bureau: International Database for fertility figures.
[25] UN has standardized data definitions and provides guidelines for organizing social and demographic surveys. (UN, 2013)

2.6.2 Analysis of Micro Data

A natural point to start is to look at total number of children for the two groups. Table C-10 presents descriptive information on average number of children by headscarf use. Women wearing the headscarf have more children on average. Even among younger cohort women born after or in 1973, still women wearing headscarf have more children on average.

We also calculated the number of children born in each year from 1990 to 2008 using NFHS-2008. Figure 2-10 depicts the average birth rate for the two groups in the period 1990-2008 whereas Figure 2-11 shows the same relationship for younger cohorts.

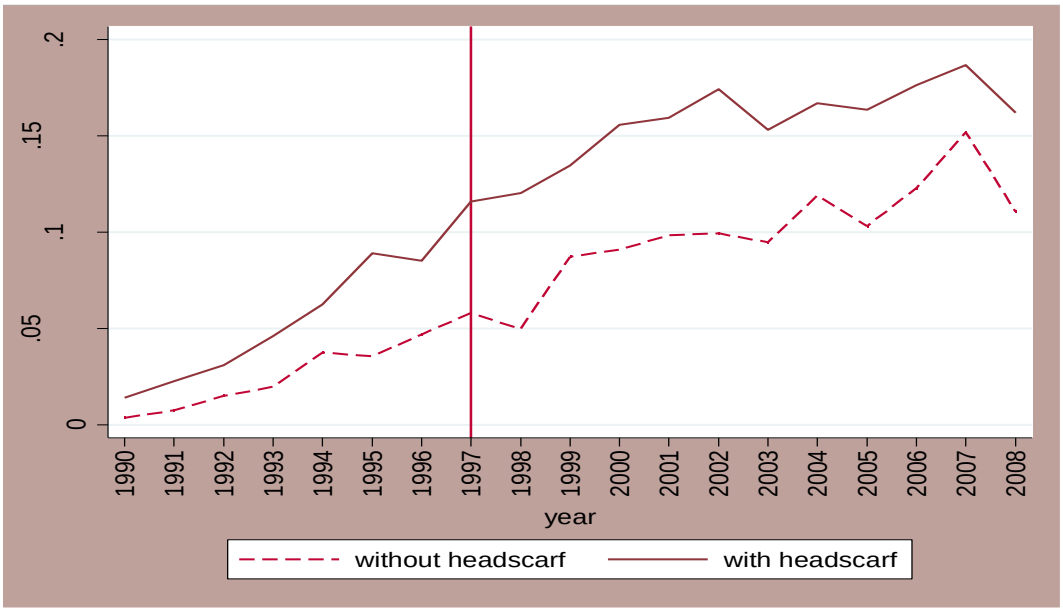
Figure 2-10: Birth Rate over Time by Headscarf Use



Source: NFHS- 2008

It is apparent from Figure 2-10 that the birth rate of women wearing the headscarf is higher than that of women not wearing it. On the other hand, both series are moving together over time. That is, there is no discernible change in birth rate of women wearing the headscarf compared to women not wearing it. Similar to tertiary education and labor force participation decisions, the headscarf ban might be more relevant for younger cohorts's childbearing. Figure 2-11 shows a drop in birth rate for women not using the headscarf, whereas we see a rise in the birth rate for women using the headscarf in 1998, one year after the ban.

Figure 2-11: Birth Rate over Time by Headscarf Use for Younger Cohorts



Source: NFHS- 2008

The difference in birth rate might be due to various factors. For instance, women not wearing a headscarf might be younger and women wearing it might be older. To control for other characteristics that might be important for childbearing, we conducted a regression analysis. In model (1), birth rate is regressed on employment status, age, marital status, region of residence, urbanity, headscarf dummy variable, after 1997 dummy variable, interaction of headscarf dummy variable with after 1997 dummy variable. This is the baseline model. The time trend is included in model (2) to capture trending. For younger cohorts, we restricted the sample only women born in or after 1973. Similar models are run, initially with the baseline regressors. This corresponds to model (3) and we add time trend in model (4). Table 2-14 presents the marginal effects calculated from the probit models for childbearing.

Table 2-14: Probit Estimates of Child Birth (Marginal Effects) – NFHS 2008

	All women		Younger Cohorts	
	(1)	(2)	(3)	(4)
Headscarf	0.011***	0.011***	-0.000	-0.000
	(3.13)	(3.12)	(-0.01)	(-0.06)
Headscarf*the ban	0.000	-0.001	0.014*	0.014*
	(0.01)	(-0.13)	(1.78)	(1.77)
Baseline Controls	+	+	+	+
Time Trend	-	+	-	+
N	138,933	138,933	77,747	77,747
Pseudo R-sqr	0.22	0.22	0.28	0.28

When the first two columns are examined, we did not observe any significant change in child birth for women wearing headscarves after 1997. On the contrary, for younger cohorts, women wearing headscarves are 1.4% more likely to have a child birth after 1997 according to model (4).

As we did for tertiary educational attainment, we check whether we observe different short-term and long term fertility behavior on different age groups. The dependent variable is number of births in the previous five years from the time of the survey. For studying the effects on short-term fertility behavior, we use NFHS-2003 dataset and for studying the effects on long-term fertility behavior, we use NFHS-2008 dataset.

For We use "born after 1980" as a cut-off point, because women born after 1980 are exposed to the headscarf ban fully whereas women before 1980 might be partially subject to the ban or might not be subject to the ban at all. Since, fertility decisions of older generations might not be comparable with younger people, women born before 1973 are also discarded from analysis.

Mother's and father's education level, region lived, type of place of residence, age, mother tongue, wealth index, household assets such as car, motorbike, TV, refrigerator, telephone as a proxy for wealth, survey year, born after 1980, headscarf use and interaction of born after 1980 and headscarf are used as baseline controls in model (1). In model (2), we differentiate between different possible effects on different age groups. The idea is that women born before 1976, if they did not repeat any grade, could get a degree without being subject to the headscarf ban, whereas women born between 1979 and 1976 were possibly at school when the ban was enacted. Therefore, we include born after 1980 and born between 1976 and 79 and their interactions with headscarf use in model (2). In model (3), we include women born in 1980 into the transition group since they might have already incurred the costs of preparing for university entrance examination. The marginal effects calculated from tobit models are shown in Table 2-6.

From model (1), women wearing a headscarf are 19.8% more likely to have birth in the previous five years from NFHS-2003. Model (2) and Model (3) also show similar results. However, according to the first three models, the interaction term is insignificant; therefore, we could not find any statistically significant effect on short-term fertility. On the other hand, according to the last three models, we do observe statistically significantly higher childbirth in the previous five years from

NFHS-2008among women wearing headscarves and born after 1980. Therefore, we conclude that the ban impacted long-term fertility among women wearing headscarves who are fully exposed to the ban.

Table 2-15: Tobit Estimates of Births in the previous five years (Marginal Effects) - NFHS

	Short-term effects			Long-term effects		
	(1)	(2)	(3)	(4)	(5)	(6)
Headscarf	0.198***	0.214***	0.218***	-0.027	-0.010	-0.003
	(4.35)	(3.32)	(3.37)	(-0.60)	(-0.15)	(-0.05)
Headscarf*Born after 1980	0.083	0.072		0.187***	0.172**	
	(1.05)	(0.79)		(3.19)	(2.25)	
Headscarf*Born between 1976&79		-0.028			-0.037	
		(-0.33)			(-0.44)	
Headscarf*Born after 1981			0.088			0.157**
			(0.89)			(1.98)
Headscarf*Born between 1976&80			-0.019			0.021
			(-0.24)			(0.26)
Baseline Controls	+	+	+	+	+	+
N	2,867	2,867	2,867	3,757	3,757	3,757
Pseudo R-sqr	0.040	0.048	0.048	0.047	0.058	0.054

2.7 Conclusion

The headscarf ban is a hotly debated issue for Turkey, and also for other countries. Although public opinion in Turkey is overall against the ban, the ban has remained since 1997. Other studies have established that the majority of females in Turkey wear headscarves. Naturally, one can expect to observe effects of the ban in many aspects of life for females, and possibly also for males. In this study, we focus on measuring the impact the headscarf ban on female educational attainment, labor force participation (LFP) and child bearing decisions.

We study the impact of the headscarf ban by employing two methodologies. We analyzed aggregate country level data using difference in differences (DD) methodology, with Turkey as the treated unit and some neighboring countries as the control group. Similarly, aggregate national data is analyzed with females as the treated and males as the control group. We also utilized individual level data to check potential effects on specific population groups. Ideally, we would need individual level data which has information on headscarf use status, educational attainment, employment status and childbearing before and after the ban. Unfortunately, we only have individual level data for headscarf use status after the ban. We can still provide suggestive evidence, because women who are born after 1980 are fully exposed to the ban, whereas women born before 1976 might not be exposed to the ban

fully if they did not repeat any grade. However, we can only observe headscarf use status once, at the time of the survey. Our analysis is based on the assumption that women's religious preferences do not change over time.

The results from country level analysis using difference in differences methodology suggest that the headscarf ban led to a 27% drop in the female to male ratio for tertiary education students, but when country specific time trends are added, the effect is no longer statistically significant. Similarly, although we find 22% drop in female LFP, when country specific time trends are added, the effect is no longer statistically significant. However, we find 0.27 increase in total fertility rate from country level analysis which includes country specific time trends.

The results from national aggregate data using males as control group also did not report any significant effect on overall female tertiary education indicators. We observe 3% drop urban LFP rate of females and 2% drop in LFP of higher educated females after the introduction of the ban compared to males. Both estimates are statistically significant at 10% level.

Although all individual level data indicate a large educational gap between women wearing headscarves and women not wearing headscarves, we did not detect a significant difference for tertiary educational attainment of women who were fully exposed to the ban (wearing headscarves and born after 1980) compared to women who may not be exposed to the ban (wearing headscarves and born between 1973 and 79) assuming standard progression through school. We also documented a wide gap in employment status of women by their use of the headscarf. Even after controlling for religion-related covariates, the use of headscarf is negatively associated with being employed. Using employment history from NFHS-2008, we find that after the enactment of the headscarf ban, employment probability dropped by 4.8% for women wearing headscarves. This is more pronounced for younger cohort women wearing headscarves. We observe a drop of 5.9% for this group after the enactment of the ban. Moreover, using full fertility, employment, marriage, and migration history from NFHS-2008, we observe an increased childbearing probability for younger cohort women wearing headscarves by 1.4% after the enactment of the ban, although the coefficient is significant at 10% significance level. We support this effect on fertility by looking at childbearing in the previous five years from the time of NFHS-2003 and 2008. Using NFHS-2003, we do not find statistically

significant effect on short term fertility of women who are fully exposed (wearing headscarves and born after 1980). But, we did find statistically significant effect on long-term fertility on women who are fully exposed.

It is plausible not to observe any significant effect on total number of tertiary education students due to excess demand for tertiary education. Those women not wearing a headscarf could easily replace women wearing scarves. Although the total number of students who get tertiary education might stay the same due to excess demand, quality might fall. In a further study, we would like to check the national placement examination scores before and after the ban to check whether there is any effect on the quality of students entering higher education institutions. It is also possible that we could not identify the impact on women wearing headscarves because, as mentioned in Cindoglu (2010), even before 1997, entering a university with a headscarf was not problem-free.

In a nutshell, we did not find any statistically significant effect of the ban on female tertiary educational attainment indicators; whereas we did find some effect at the national and individual level on female labor force participation indicators and fertility. Even if one does not care of the consequences of the ban for this particular group of women, such a ban could have an unintended byproduct of an increased number of people raised by mothers who prefer to use headscarf. Therefore, we suggest policy makers to consider the potential effects of banning headscarves on women's educational attainment, employment opportunities and fertility when addressing concerns of secularism.

Appendix A: Data Description

A.1. Aggregate National Tertiary Education Indicators

National female labor force participation figures are obtained from the TURKSTAT's webpage under Population, Demography, Housing & Gender main tab ➔ Gender, Life And Family ➔ Data ➔ Gender Indicators ➔ Labour Force ➔ Labour force by household population. These statistics are gathered from "Household Labor Force Surveys".

Public sector employment figures are obtained from the TURKSTAT's webpage under Population, Demography, Housing & Gender main tab ➔ Gender, Life and Family ➔ Data ➔ Gender Indicators ➔ Labour Force ➔ Employment by status of workplace. These statistics are gathered from "Household Labor Force Surveys".

LFP rate of Higher Educated women, men and their unemployment rate are obtained from the TURKSTAT's webpage, under Employment, Unemployment & Wages tab ➔ Labor Force Statistics➔ Data ➔ Statistical Tables ➔ Periodic Results Of Household Labour Force Survey ➔ Turkey ➔ Labour Force Status By Educational Level. These statistics are gathered from "Household Labor Force Surveys".

Urban female LFP and female urban non-agricultural unemployment rate are also obtained from TURKSTAT's webpage, under Employment, Unemployment & Wages tab ➔ Labor Force Statistics➔ Data ➔ Statistical Tables ➔ Periodic Results Of Household Labour Force Survey ➔ Urban ➔ Labour Force Status By Non-Institutional Population, Years And Sex. These statistics are also gathered from "Household Labor Force Surveys."

A.2. Country Comparison Data

Data on the ratio of females to males in tertiary education students in Syria from 1990 to 1995 and that of Turkey from 1990 to 2007 except for 1996 were also taken from the World Bank Dataset from http://data.worldbank.org/indicator/SE.ENR.TERT.FM.ZS webpage. The figures for Syria from 1998 to 2007 are from Statistical Institute of Syria's webpage under Statistical Abstracts. Figures from 1998 to 2002 are from Statistical Abstract, 2003 and figures from 2003 to 2007 are from Statistical Abstract, 2008. The figure for Turkey in 1996 is calculated used Higher Education Statistics Yearbook of Turkey since it was missing in World Bank Dataset.

LFP figures for Iran, Syria, Egypt and Turkey are downloaded from the World Bank's Dataset from http://data.worldbank.org/indicator/SL.TLF.CACT.FE.ZS webpage. We used World Bank data for Turkey to avoid incomparability.

Ratio of Female LFP to Male LFP figures for Iran, Syria, Egypt and Turkey are downloaded from the World Bank's Dataset from http://data.worldbank.org/indicator/SL.TLF.CACT.FM.ZS webpage. We used World Bank data for Turkey to avoid incomparability.

Total fertility rate (births per women) for Iran, Syria, Egypt and Turkey are downloaded from the World Bank's Dataset from http://data.worldbank.org/indicator/SP.DYN.TFRT.IN webpage.

A.3. Micro Data

Firstly, we use data from "Religion, Secularism and Headscarf in Daily Life Survey" (Gündelik Yaşamda Din, Laiklik ve Türban Araştırması) to get information on educational attainment and labor market outcomes of women wearing a headscarf. Religion, Secularism and Headscarf in Daily Life Survey (RSHDLS) is a cross-section survey conducted in September, 2007 by Konda Research Company. Konda Research Company is a public opinion poll company which specializes in collecting and analyzing data for both political and sociological quests. This survey provides data on various indicators of religious practices, perceptions on secularism and the use of headscarves. Data is collected through face to face interviews in all regions of Turkey. Individuals eighteen and older are surveyed, and the sample is representative of Turkey's non-institutionalized adult population. The sample size is 5,291. We restricted our sample to 2,639 females to study the effect on women. The wording of the headscarf question is as follows: "Do you wear headscarves while going out? If so, how?" Possible answer categories are; Do not wear headscarves, Headscarves (basortusu) and a more traditional headcover (Yemeni), A way of calling headscarves which become popular in last decade (Turban), Full-face veil (carsaf veya pece). 67% of women reported wearing some sort of headscarf. 75% of women who wear some sort of headcover describe it as a headscarf. For our study, we describe women as wearing headscarves if they report wearing one of three types.

Secondly, we use two different surveys conducted by A&G Research Company. A&G Research Company specializes in providing services to political parties, and is well known for its success in predicting election outcomes. The first survey was conducted in 2003 for a well-known Turkish newspaper called "Milliyet," with 1,881 respondents; the second one was conducted in 2007 for a famous TV channel called "Kanal D," with 2,620 respondents. These surveys are mainly conducted to collect data on political

preferences, opinion of the public on the headscarf and to measure how the use of headscarf and other religious practices are changing in society. Data is collected by face to face interviews in all regions of Turkey. Individuals eighteen and older are surveyed and the samples are representative of Turkey's electorate adult population. For the purpose of this study, we pooled both datasets together, which provide us with 4,501 observations. We discard males, thus, our analysis is based on 2,245 females. The wording of the headscarf question is the same for the two surveys, which is as follows: "Is there anyone at this household who covers her head when she goes out for shopping, city center, walking etc. If so, who? The answers are categorized as: Yes, I do, Yes, my daughter, Yes, my mother, Yes, my grandmother, Yes, others. No. For our study, we describe women as wearing headscarves if they answer as "Yes, I do". 66.24% of women reported wearing a headscarf in the 2003 survey, whereas this percentage has dropped to 57.45% in 2007.

Lastly, we use two rounds of National Family and Health Survey (NFHS). The NFHS is mainly designed to assess women's and their kids' health and nutrition status, fertility history, health related knowledge, and women's status in society. It is administered to married women between the ages of 15 to 49. In the last two waves, some religious behavior is also covered in the survey. For the purpose of this study, we pooled together the NFHS-3 conducted in 2003 and the latest NFHS-4 conducted in 2008. The data sets report information on the age, sex, health, education and employment status of the individual, as well as information on religious behavior such as the use of a headscarf, frequency of five-daily prayers, and fasting. NFHS-3 has a sample size of 8,075 women and NFHS-4 has a sample size of 7,405 women. Thus, we have in total 15,480 women in our sample. The wording of the headscarf question is the as follows: "Do you wear a head scarf when you go outside the street?" The answers are categorized as; Yes and No in 2003. In 2008, for the same question, the answers are categorized as; Yes, regularly, yes, irregularly and No. 75.4% and 75.6% of women reported wearing headscarf in 2003 and 2008 survey respectively.

Appendix B:

Table B-1: Religiosity Changes over Time among Women

	1990	1996	2001	2007	Average
A religious person	79.8	78.46	82.33	85	81.68
Not a religious person	19.6	21.26	16.95	14.7	17.78
A convinced atheist	0.59	0.28	0.72	0.3	0.54
Total Number of Respondents	505	715	1,658	660	3,538

Table B-2: Probit Estimates of Headscarf Status (Marginal Effects) - NFHS-2003&2008

	(1)	(2)	(3)	(4)
Born after 1980	0.035*	0.024		
	(1.79)	(0.73)		
Born between 1976&79		-0.008		
		(-0.43)		
Headscarf*Born after 1981			0.023	0.006
			(1.21)	(0.18)
Born between 1976&80				-0.010
				(-0.54)
Baseline Controls	+	+	+	+
N	6,118	6,118	6,118	6,118
Pseudo R-sqr	0.187	0.187	0.187	0.187

Baseline controls: survey year, age, wealth index, mother's education level, father's education level, ownership of car, motorbike, TV, refrigerator and telephone as a proxy for wealth, region lived, type of place of residence, mother tongue.

Appendix C: Descriptive Statistics

Table C-1: Descriptive Statistics for Ratio of Females to Males in Tertiary Education Students

	Before 1997	After 1997	Difference	DD
Turkey	57.83	71.84	15.21	-23.06
Iran	43.51	99.7	56.17	
Syria	68.1	86.56	18.49	

Table C-2: Women's education level by use of headscarf (N=2,609 women in total)

	without headscarf	with headscarf	P-value
Illiterate	2.09	17.39	0.00
literate without a diploma	0.58	7.61	0.00
primary school graduate	25.09	50.8	0.00
middle school graduate	12.31	10.93	0.30
high school graduate	42.28	12.13	0.00
university or higher degree	17.65	1.14	0.00

Source: Konda Dataset, 2007

Table C-3: Women's education level by use of headscarf (N=2,241 women in total)

	without headscarf	with headscarf	P-value
no education	2.87	18.39	0.00
literate without a diploma	1.61	6.06	0.00
primary school graduate	21.47	52.85	0.00
middle school graduate	9.87	10.73	0.52
high school graduate	44.32	10.15	0.00
university or higher degree	19.86	1.82	0.00

Source: A&G Research Company 2003& 2007 combined

Table C-4: Women's education level by use of headscarf (N=15,456 women in total)

	without headscarf	with headscarf	P-value
no education	2.97	23.78	0.00
incomplete primary education	2.25	6.84	0.00
primary school graduate	31.41	53.2	0.00
incomplete secondary education	17.19	7.94	0.00
secondary school graduate	26.46	6.24	0.00
University or higher education	19.73	2	0.00

Source: NFHS-3 in 2003 & NFHS-4 in 2008

Table C-5: Descriptive Statistics: Female Labor Force Participation

	Before 1997	After 1997	Difference	DD
Turkey	31.76	26.43	-5.32	-5.6
Iran	10.13	15.59	5.46	
Syria	20.76	18.34	-2.42	
Egypt	22.59	23.41	-2.23	

Table C-6: Women's Labor Market Status by headscarf - Konda 2007 (N=2,616)

	without headscarf	with headscarf	P-value
Employed	**30.32**	**10.22**	**0.00**
Public Sector	7.06	0.57	0.00
Private Paid Employment	13.66	5.14	0.00
Self-Employment and Other Jobs	9.61	4.51	0.00
Not Employed	**69.68**	**89.78**	**0.00**
Unemployed	5.90	1.48	0.00
Student	11.23	1.66	0.00
Not in Labor Force	52.55	86.64	0.00

Table C-7: Women's Labor Market Status by headscarf A&G 2003& 2007 (N=2,233)

	without headscarf	with headscarf	P-value
Employed	**31.39**	**5.81**	**0.00**
Public Sector	8.93	0.88	0.00
Private Paid Employment	12.49	2.72	0.00
Self-Employment	9.97	2.21	0.00
Not Employed	**68.61**	**94.19**	**0.00**
Unemployed	2.41	0.37	0.00
Not in the Labor Force	50.63	92.87	0.00
Student	15.58	0.96	0.00

Table C-8: Women's Labor Market Status by headscarf NFHS- 2003&2008 (N=15,455)

	without headscarf	with headscarf	P-value
Employed	33.99	25.39	0.00
Not Employed	66.01	74.61	0.00

Table C-9: Descriptive Statistics - Total Fertility Rate

	Before 1997	After 1997	Difference	DD
Turkey	2.86	2.3	-0.56	0.71
Iran	3.84	2.07	-1.78	
Syria	4.73	3.46	-1.27	
Egypt	3.92	3.16	-0.76	

Table C-10: Average number of children by use of headscarf NFHS - 2008

	without headscarf	with headscarf	P-value
All women (N= 7,390)	1.7	2.96	0.00
Younger cohorts (N= 4,122)	1.36	2.23	0.00

Appendix D: Summary Table

Table D- Summary of Regression Results				
Micro Data				
		Konda	A&G	NFHS
Tertiary Degree Holding	Headscarf	-0.081***	-0.143***	-0.053***
	Headscarf*age at most 17 in 19997		-0.005	
	Headscarf*born after 1980			-0.001
Employment	Headscarf	-0.041**	-0.049*	-0.031**
	Headscarf*the ban			-0.048***
Childbearing for younger cohort	Headscarf*the ban			0.014*
Aggregate Data				
Log of F/M ratio in Tertiary Stud.	-0.04			
Log of LFP	-0.01			
Total Fertility Rate	0.027**			

3 Does presumed consent save lives? Evidence from Europe

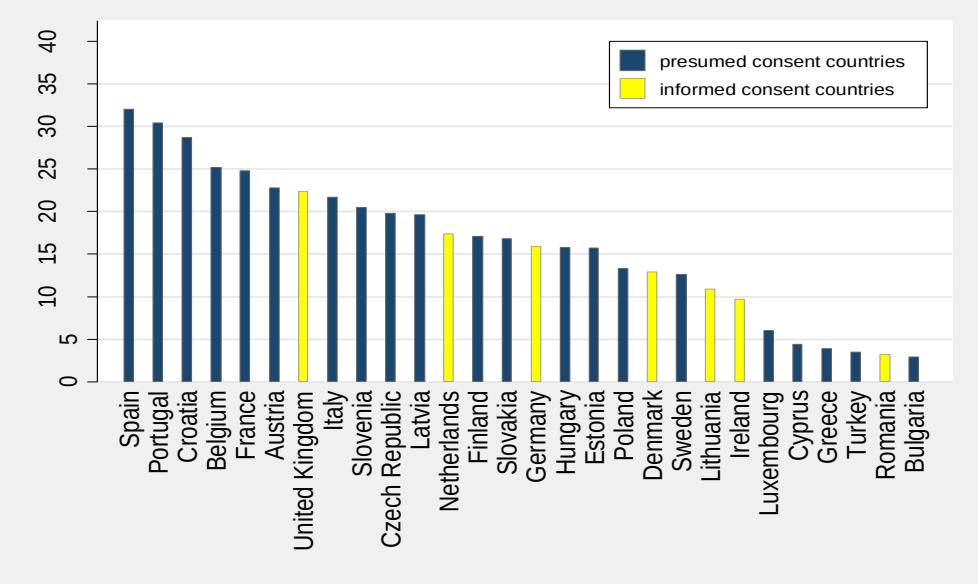
3.1 Introduction

The number of organ transplants has constantly increased in each year in Europe; but has not kept pace with demand. The long waiting lists for transplantation became a common phenomenon. Among others, one policy option is playing with legislative defaults for fostering cadaveric donation.

In Europe, there are two types of institutional arrangements for getting consent for organ transplantation. One is informed consent or opt-in and the other is presumed consent or opt-out. In informed consent regime, individuals are expected to declare explicitly their willingness by registering to be an organ donor. Therefore, individuals who are not registered in the system are assumed to not donate their organs in the case of death. In presumed consent regime, a brain-dead individual whose organs are suitable for transplantation is automatically considered to be a donor unless she has stated a preference for not donating. In practice, in some presumed and informed consent countries, consent from the family of the deceased is routinely sought even if he or she explicitly stated her preference to be a donor.

Figure 3-1 shows cadaveric donor rate in 2010 for EU-27 countries, Croatia and Turkey by the type of consent regime. It is apparent from Figure 3-1 that cadaveric donor rate is higher in presumed consent countries. Is this by chance?

Figure 3-1: Cadaveric Organ Donor Rate per million population in 2010



Source: The International Registry of Organ Donation and Transplantation (IRODaT)

Theoretically speaking, if the costs associated with registering preferences for organ donation are low[26], the defaults would not have a great effect for fully rational individuals who already have established preferences for organ donation. This is because in the case of mismatch of the default and the preferences, individuals are expected to take an action for the desired option. On the other hand, if individuals are more likely to accept the effortless default option rather than make a choice which has mental costs, particularly for organ donation since it requires thinking about death which is generally perceived as unpleasant and stressful, then defaults might matter.

Even if presumed consent is not strictly enforced, that is, family consent is always sought; legislating presumed consent might still be a positive signal from the government to the families. Both informed and presumed consent legislation carry a message about a social norm on the default course of action. The signal in presumed consent is that the government expects the family to give consent; whereas in informed consent legislation, donating is framed as something extraordinary, up to the wishes of the family.

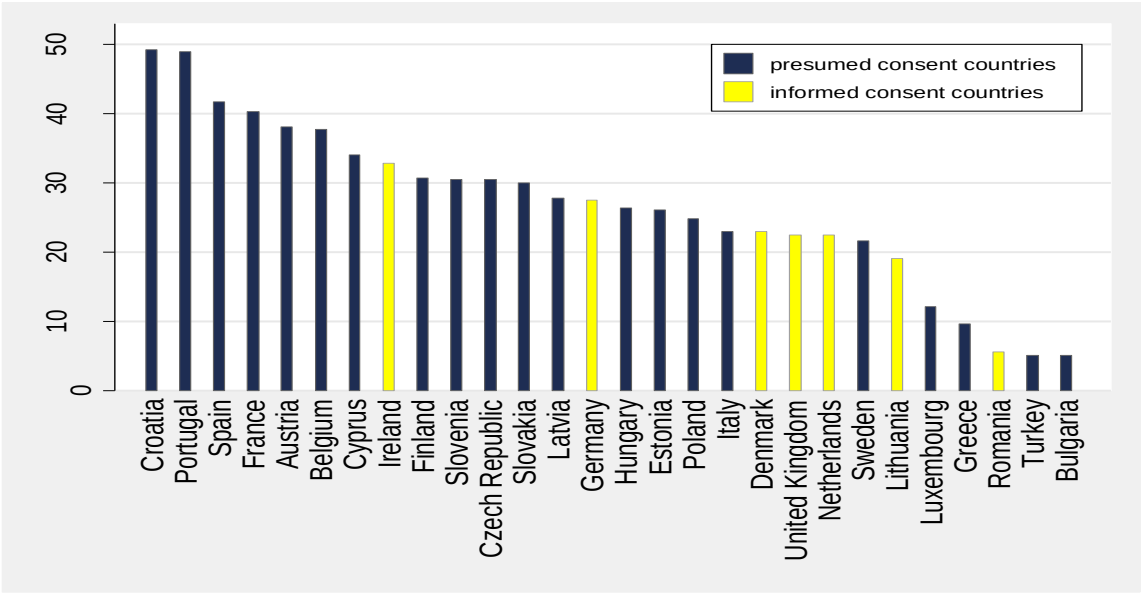
Presumed consent legislation might even impact the way doctors talk to families. Doctors have the law on their side when trying to explain the need for organ donation to a bereaved family, whereas in informed consent countries doctors' task of convincing a bereaved family is much more difficult and there is not much to say in case a family decides not to donate organs.

The underlying motivation for studying legislative defaults is not increasing organ donors per se, what matters from a policy perspective is the number of "lives saved". It is not clear whether increased cadaveric donation effectively translates into number of transplantations. Moreover, even if informed consent countries succeed in producing the same number of cadaveric donor rate pmp (per million population), there might still be differences in terms of age structure of the donor pool. A person's organs could only be used if she/he dies through unexpected causes of death such as homicide, transport accidents and cerebro-vascular diseases. Especially young people who have a riskier life-style are more likely to end up dying through those causes. However, these people are also

[26] There are possibly differences in registration costs among countries; some countries allow for registering the preferences for organ donation through online forms, some others have special telephone lines. Some only accept applications through mails. Some only accept registration at some special locations which might require incurring transportation costs.

the ones who are less likely to think of death and register as organ donors since they may foresee a long life-time. Presumed consent could solve the registration problem of young and risk-taking people. On the other hand, elderly are more likely to think of death and their preferences for organ donation might be better known by their close relatives. However, their organs are less useful compared to those people who die at younger ages. Therefore, we expected to observe higher kidney transplantation rates in presumed consent countries since presumed consent is supposed to solve procrastination behavior of young people who are least likely to think of death and whose organs are the most valuable.

Figure 3-2: Kidney Transplantation from Deceased Donors Rate per Million Population in 2010



Source: IRODaT

Figure 3-2 shows kidney transplantations from deceased donors in 2010 for the same set of countries by their legislative setting. We also observe higher kidney transplantation rate in presumed consent countries compared to informed consent countries.

This study aims at examining the impact of presumed consent on cadaveric donations and kidney transplantations using a panel dataset from the EU-27 countries plus Croatia in the period 2000-2010. For identifying the impact of presumed consent, we would ideally need country fixed effects models which would treat unobserved country level heterogeneity. However, there are a few changes in legislation over the last 20 years in Europe. Therefore, we cannot estimate country fixed effects reliably and bound to use pooled OLS estimates. Yet, pooled OLS analysis would be biased if presumed consent is legislated in countries where there is higher social acceptance of organ donation.

Therefore, in this study we consider the relationship between presumed consent and several organ donation indicators. These indicators are willingness to donate one's own organs, willingness to give consent for a family member, organ donation card holding, actual cadaveric donation rates, and kidney transplantation rates. To ensure reliability of pooled OLS estimates, we follow a three step approach.

We firstly study differences in willingness to donate one's organs in presumed and informed consent countries. If we do not find any statistically significant difference in willingness to donate one's organs in presumed and informed consent countries, there will be less concern for biased estimates from pooled OLS analysis. Secondly, we study differences in registering preferences for organ donation in presumed and informed consent countries by looking at organ donation card holding behavior. For presumed consent to have an impact on organ donation rates, we should observe differences in registering behavior. If people do register their preferences for organ donation in case of a mismatch between their preferences for organ donation and legislative default, then it is unlikely to observe any behavioral effects of presumed consent. The third step which forms our main analysis explores the impact of presumed consent legislation on cadaveric donations and kidney transplantations.

In the first step, we check endogeneity of presumed consent by asking the following question "Do individuals in presumed consent countries exhibit higher willingness to donate their own organs and higher willingness to give consent for a family member?" For answering this question, individual level data from the 2002, 2006, 2009 Eurobarometer Surveys are used. Even after controlling for socio-economic background indicators, we do not detect any statistically significant relationship between willingness to donate and presumed consent legislation. These findings imply that presumed consent legislation is not necessarily enacted in countries where there is wide social acceptance of organ donation. This is somewhat reassuring for Pooled OLS results.

In the second step, we try to answer the following question "Do individuals take action in line with their preferences when there is a mismatch of the legislative default and the desired option?" To check on this idea, we use organ donation card holding from Eurobarometer, 2006 survey since having it requires registering to the organ donation authority of a country. In line with our

expectations, we find significantly lower donation card holding among those who are willing to donate their organs in presumed consent countries. Surprisingly, among people who are not willing to donate, we do not observe higher registration in presumed consent countries. These findings suggests that presumed consent can increase cadaveric donation rates because people who are not willing to donate their organs fail to register their preferences in presumed consent countries.

In our main analysis, we try to answer the following questions: Does presumed consent impact cadaveric organ donations? Does presumed consent increase kidney transplantations? Using international organ donation registry data, we find that presumed consent countries have 28 to 32% higher cadaveric donation and 27 to 31% higher kidney transplant rates in comparison to informed consent countries after accounting for potential confounding factors.

Evidence from other studies such as Madrian and Shea (2001), Choi, Laibson and Madrian (2004) and Johnson, et al. (1993) show that cleverly set defaults affect the pension savings and insurance choices. In the context of organ donation, a handful of papers has analyzed whether legislative defaults such as presumed consent would result in more deceased organ donation rates (Johnson and Goldstein 2003; Abadie and Gay 2005; Healy 2005; Bilgel 2010)

Although previous studies found higher cadaveric organ donation rates in the presumed consent countries compared to informed consent countries, there is no consensus about the underlying mechanism. Some researchers (the first group) attribute higher cadaveric organ donation rates to the effect of presumed consent legislation, whereas others (the second group) see the presumed consent legislation as an indicator of a country's commitment to organ donation. Mainly, the first group of studies is criticized on the ground that they did not address unobserved heterogeneity adequately. This unobserved heterogeneity could be dealt with a country fixed effects model. However, there is either very little change in the legislation over time. Therefore, fixed effects models could not be run. Pooled OLS analysis would be biased if presumed consent is enacted in countries where there is higher social acceptance of organ donation because, the coefficient of presumed consent might capture the effect of social acceptance of organ donation.

This study contributes to the literature in some important ways. We firstly address potential endogeneity of presumed consent by showing evidence that presumed consent is not necessarily

legislated in countries where there is higher social acceptance of organ donation. To address unobserved heterogeneity even better, we group countries according to their geographic, ethnic, cultural, and organ donation related cooperation. We then identify the impact of presumed consent running country group fixed effects models. The results still show higher cadaveric donation rates in presumed consent counties which suggest that accounting for unobserved heterogeneity ultimately appears to have little effect.

Secondly, we address the claim that presumed consent is an indicator of a country's commitment to organ donation rather than a causal mechanism in itself. We show that after taking into account a country's commitment to organ donation proxied by kidney transplant centers as an additional control variable, the coefficient of presumed consent is still statistically significant and it even increases. Thirdly, although according to Eurobarometer, 2009 survey on organ donation, religious reasons, distrust in the system and scare of manipulation of the human body are three major causes of refusal for organ donation, previous studies have not dealt with trust in the system and religion differences adequately. Abadie and Gay (2005) include religion with a Catholic country indicator which is based on majority of population being Catholic or not. To capture trust in the system, we included corruption perceptions scores from Transparency International. To control for religiosity changes over time, we compiled percentage of population being Roman Catholic and having no religion mainly from International Social Survey Program (ISSP), European Social Survey (ESS), European Values Survey (EVS) and Eurobarometers conducted between 1999/2000 and 2010. Lastly, to the author's knowledge, this study is the first to analyze the impact of presumed consent on kidney transplantation which is more relevant from a policy perspective.

3.2 Organ Shortage Problem and Potential Solutions

In this section, we show the extent of the organ shortage problem and discuss proposals suggested to minimize it. Table 3-1 shows some transplantation (TX) indicators for EU-27+ Croatia. In 2008, 34,003 patients are added to the waiting list and only 9,042 cadaveric organ donations were available for transplantation for EU-27 countries plus Croatia. The rest of the transplantations were conducted using organs from living donors. As of 2008, over 63,000 people were awaiting one of the

aforementioned transplant procedures and this figure has increased to 64,663 patients in year 2009. The problem is not only the waiting time but also 3,772 patients died while waiting for transplants in 2008 and 3,809 people lost their lives while waiting for an organ in 2009 for the same set of countries.

Table 3-1: Transplantation Indicators in EU-27+Crotia		
	2009	2008
Total Number of Kidney Transplants	18,056	17,356
# of Patients awaiting for a TX by 31st Dec	64,663	63,490
# of Patients died while on the Waiting list	3,809	3,772
# of Cadaveric Organ Donors[27]	9,230	9,042

Source: Newsletter Transplant (2010; 2009)

To alleviate organ shortage, increasing live and/or cadaveric donation are possible solutions. One the one hand, living donation seems promising since living donation provides a better outcome of patient survival when compared with deceased-donor transplantation (Davis & Delmonico, 2005). However, live donation is possible only for certain organs such as kidneys, part of liver and lung. Heart could only be obtained from cadaveric donors. Moreover, Ellison et al. (2002) identified 56 live kidney donors in the US who were subsequently listed for a kidney transplant themselves. Furthermore, cadaveric organ donation still provides the majority of organs. Therefore, in this study, we chose to focus on cadaveric organ donors.

Cadaveric donor rate may be influenced by educational efforts of governments, public awareness campaigns of the bodies responsible for organ donation, religious setting and transplant infrastructure together with the type of consent regime prevailing in the countries. However, educational efforts and public awareness campaigns do not always translate into an increased number of donors. For example, in the Netherlands, in 1998, the government organized sending 12 million letters in a country of 16 million asking citizens to register, which failed to impact the effective consent rate (Oz et al, 2003). The ineffectiveness of the campaign also casts doubt on the role of education since the Netherlands is a country with a highly educated population.

Having observed demand for organs exceeding supply of organs, some economists emphasized the use of monetary incentives for increasing the supply of organs from both cadaveric donors as well as live donors (Cohen 1989; Becker and Elías 2007; Howard 2007). However, the use of financial

[27] Includes non-heart beating organ donors.

incentives is a controversial suggestion since transactions on human body are often considered to be repugnant (Roth 2007). Oz et al. (2003) reports that 66% of international transplant medical professionals oppose to direct compensation methods such as tax credits or life insurance benefits for donors.

According to international transplantation medical professionals, the most effective option for alleviating organ shortage is enactment of presumed consent. (Oz, et al., 2003) The other potential channels such as increasing awareness through campaigns and education do provide nudges. However, these nudges do not move the masses. Moreover, these nudges require investment which has a very low probability of return because the likelihood of dying under conditions which would be suitable organs to be transplanted is very low. Howard and Byrne (2007) estimate the probability of a potential donor being an actual donor at some point in her lifetime is 0.0028. Therefore, investing sizeable amounts of money for convincing people for organ donation might produce a very negligible effect on organ donation. So, we choose to study the potential of the presumed consent legislation for producing more organ donors.

There are two strands of literature which analyzes the impact of presumed consent on organ donation. The first strand studies organ donation rates before and after the enactment of presumed consent legislation. The second strand compares organ donation rates in presumed consent countries with respect to informed consent countries.

From the first strand, all studies report higher organ donation rates after the enactment of the presumed consent legislation in Austria and Belgium. After the introduction of presumed consent in Austria in 1982, the donor rate has quadrupled. (Gnant et al, 1991) Similarly, Roels et al. (1991) report more than doubling of kidney donation in Belgium after the introduction of presumed consent in 1986. I cannot utilize the legislation change in Austria and Belgium since these changes occurred before international country-level organ donation data is available[28].

The second strand also reports higher donation rates when presumed consent countries are compared with informed consent countries. Johnson and Goldstein (2003) studied 10 European countries national registries for organ donation and find evidence higher effective consent rate in

[28] The IRODaT provides organ donation rates from 1993 and transplantation rates from 1999.

presumed consent countries. They also find the relevance of presumed consent using online experiment. Abadie and Gay (2005) studies 22 Western developed countries and find higher organ donation rates in presumed consent countries for 1993-2002 period. However, since only Sweden changed its legislation from informed consent to presumed consent in 1996, that studies relied on pooled cross-section analysis in which they did not take into account country fixed effects. For instance, if presumed consent is enacted in countries where there is higher social acceptance of organ donation, the found effect might be biased. Healy (2005) studied 17 OECD countries higher cadaveric donation rates and finds higher donation rates in presumed consent countries. However, Healy (2005) attributes higher donation rates in presumed consent countries to commitment of these countries for organ donation rather than a causal effect of the presumed consent legislation. Bilgel (2010) also studied organ donation for 1993-2006 period for larger number of countries. However, that study used Fixed Effects Vector Decomposition methodology which is not econometrically valid. (Greene, 2010)

3.3 Empirical Analysis

In this study, our aim is to examine the impact of presumed consent on cadaveric donations and kidney transplantations. However, there are a few changes in legislation over the last 20 years in Europe. Given that we cannot estimate country fixed effects, our pooled OLS analysis would be biased if presumed consent is legislated in countries where there is higher social acceptance of organ donation. Therefore, we follow a three step approach.

In the first step, we study differences in willingness to donate one's organs between presumed and informed consent countries. If we do not find any statistically significant difference in willingness to donate one's organs in presumed and informed consent countries, there will be less concern for pooled OLS analysis. In the second step, we study differences in registering one's preferences for organ donation between presumed and informed consent countries. For presumed consent to have an impact on organ donation rates, we should observe differences in registering behavior. If people do register their preferences for organ donation in case of a mismatch between their preferences for organ donation and legislative default, then it is unlikely to observe any behavioral effects of presumed

consent. The third step which forms our main analysis explores the impact of presumed consent legislation on the cadaveric donations and kidney transplantations.

For the purposes of this study, we use EU countries and some other countries for which Eurobarometer survey is conducted either in 2002, 2006 or in 2009. These countries/regions are Austria, Belgium, Bulgaria, Croatia, Cyprus Republic, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Ireland, Italy, Latvia, Lithuania, Luxembourg, Malta, Macedonia, Netherlands, Poland, Portugal, Romania, Slovakia, Slovenia, Spain, Sweden, Turkey, Turkish Cypriot Community, and United Kingdom. Macedonia, Turkish Cypriot Community are eliminated because organ donation figures are not available for them. Malta is discarded because type of legislation could not be found for this country. Turkey is discarded because causes of deaths which constitute the bulk of cadaveric donors were not available.

3.3.1 The Association between Willingness to Donate and Presumed Consent

The motivation for undertaking this analysis is that the effect of presumed consent in our main analysis could be biased if presumed consent proxies willingness to donate. If there is no statistically significant difference in willingness to donate one's own organs and give consent for a family member in presumed and informed consent countries, there will be less concern for pooled OLS analysis. Therefore, in this section, we analyze whether willingness to donate, willingness to give consent for a family member is higher in presumed consent countries using individual level data from 2002, 2006, 2009 Eurobarometer Surveys. Willingness to donate own organs variable is derived from the following question: "Would you be willing to donate one of your organs to an organ donation service immediately after your death?" Willingness to donate organs of a family member is derived from "If you were asked in a hospital to donate an organ from a deceased close family member, would you agree?" question.

Eurobarometer surveys are conducted with face to face interview method. Potentially, responses to these questions might not fully reflect the true preference for organ donation since the respondents' tendency to give socially desirable answers. However, this does not pose a serious concern for our

study as long as people from informed and presumed consent countries do not differ in their tendency to engage in socially desirable responding.

Table 3-2: Attitudes to Organ Donation by Consent Regime

	% Willing to Donate (own)		% Willing to Donate (Family Member)	
	Presumed	Informed	Presumed	Informed
Yes	55.83	55.56	52.18	50.81
No	26.99	25.45	25.08	23.18
Do not Know	17.18	19.00	22.73	26.01
N	48,608	22,494	48,608	22,494

Table 3-2 provides descriptive information on the willingness to donate indicators. There are not significant differences in terms of percentage of individuals who are willing to donate their own organs between presumed and informed consent countries. However, individuals who are willing to give consent for a family member are significantly more prevalent in presumed consent countries compared to informed consent countries. For both indicators, a large percentage of people did not make up their mind yet and replied as "do not know" which implies the preferences for organ donation are not clear cut. Therefore, presumed consent legislation could produce more organ donation especially through people who do not have a preference.

Using these surveys, we estimate OLS regression of willingness to donate one's own organs and willingness to give consent for a deceased relative on an indicator variable of presumed consent in model (1). We add to these regressions a set of control variables in model (2). The control variables are gender, age, residence in an urban or small town (vs. the omitted category of large town), country marital status, occupation, age of the respondent at which her full-time education has ended (as an indicator for education), fixed telephone line and mobile telephone ownership as a proxy for wealth[29].

We also experiment with different samples. Initial regression analysis is conducted for the whole sample. We excluded Spain from the sample to see whether anything changes since Spain is well-known for its success in organ donation (Matesanz and Miranda 2002; Chang et al. 2003). The regression models are estimated with standard errors clustered at country level. We also estimated

[29] We used fixed telephone line and mobile phone ownership as a proxy for wealth since this information was available in all three Eurobarometer surveys. Eurobarometer 2002 survey has household income variable for each country whereas Eurobarometer 2006 and 2009 surveys has asked ownership of a set of household durables such as television, DVD player, music CD player, computer, internet access, car etc. Collecting information on ownership of household durables became popular with DHS (Demographic Health Surveys) use of these as a proxy for wealth. Moreover, Filmer & Pritchett (2001) argue that ownership information can be used as a good proxy.

these models with ordered probit models. The results are very similar. For the ease of interpreting the regression coefficients, Pooled OLS results are displayed in Table 3-3.

Table 3-3: Pooled OLS Willingness to Donate Regression Results

	Willingness to donate (own)		Willingness to donate (Family Member)	
	(1)	(2)	(1)	(2)
a. All				
Presumed	-0.013	0.003	-0.005	0.008
	(0.085)	(0.073)	(0.069)	(0.063)
Baseline Controls	-	+	-	+
N	71,102	70,085	71,102	70,085
b. Excluding Spain				
Presumed	-0.021	-0.009	-0.016	-0.006
	(0.086)	(0.073)	(0.070)	(0.063)
Baseline Controls	-	+	-	+
N	68,073	67,085	68,073	67,085

Standard errors (in parentheses) are clustered at country level, * $p < .1$, ** $p < .05$, *** $p < .01$

Interestingly, we find negative coefficients for the relationship between willingness to donate one's own organs, willingness to give consent for a deceased family member and presumed consent legislation for the whole sample. After the control variables are accounted for, the coefficient becomes positive yet insignificant. We did not find any statistically significant relationship in any of these models. These findings imply that presumed consent legislation is not necessarily enacted in countries where there is wide social acceptance of organ donation.

3.3.2 The Association between Registering Preferences and Presumed Consent

In this section, we analyze whether individuals take action in line with their preferences for organ donation when there is a mismatch of the legislative default and their preferences. If people do register their preferences for organ donation in case of a mismatch, then it is unlikely to observe any behavioral effects of presumed consent. To check on this idea, we use organ donation card holding from Eurobarometer, 2006 survey since having the card requires registering to the organ donation authority in a country. Donation card holding variable is derived from the following question: "Do you already have an organ donation card?"

In general, we expect individuals who are willing to donate their organs to register their willingness in informed consent countries with the donation card and individuals who do not want to donate their organs to register their refusal with the donation card in presumed consent countries.

However, we see individuals from presumed consent countries register their willingness with the donation card since most of the presumed consent countries ask for family consent and family refusal is minimized when the deceased already has the donation card. (Siminoff et al, 2001)

We also observe individuals from informed consent countries to register their non-willingness with the donation card. Letting individuals to register their non-willingness might be instrumental for avoiding family refusal in informed consent countries. If a person did not register to non-willing list, family members might be more likely to give consent since this signals that the person is not strongly against organ donation.

Table 3-4 provides descriptive information on the donation card holding. Almost 6% of individuals in presumed consent countries have the donation card, whereas 19% of individuals in informed consent countries have the donation card. Among individuals who are willing to donate their organs, 32% of them registered their preferences with the donation card in informed consent countries, whereas almost 10% of individuals who are willing to donate their organs also have registered their willingness in presumed consent countries. Among individuals who are not willing to donate their organs, roughly 2% and 3% of them registered their preferences in presumed and informed consent countries respectively. In line with intuition, among individuals who are willing to donate, registering preferences with organ donation card is less common in presumed consent countries. But, surprisingly, we observe higher registration in informed consent when individuals are not willing to donate.

Table 3-4: Percentage of Organ Donation Card Holding by Consent Regime and Preferences for Organ Donation

	Total Sample	Willing to Donate	Not Willing to Donate
Presumed	5.97	9.72	1.56
Informed	18.93	32.11	3.1
N	27,584	15,053	7,793

We firstly regress having the donation card on presumed consent dummy variable in model (1) for the whole sample. We add to this regression a set of control variables in model (2). The control variables are gender, age, residence in an urban or small town (vs. the omitted category of large town), country marital status, occupation, age of the respondent at which her full-time education has ended (as an indicator for education), fixed telephone line and mobile telephone ownership as a proxy for

wealth. The same models are also run for subsample of individuals who are willing to donate or not willing to donate.

We also experiment with different samples. For instance, we excluded Spain from the sample to see whether anything changes. All regression models are estimated with standard errors clustered at country level. For the ease of interpreting the coefficients, marginal effects calculated from the probit estimation results are displayed in Table 3-5.

Table 3-5: Organ Donation Card Holding Regression Results

	Total Sample		Willing to Donate		Not Willing to Donate	
	(1)	(2)	(1)	(2)	(1)	(2)
a. All countries						
Presumed	-0.130**	-0.118**	-0.224***	-0.214***	-0.015	-0.012
	(0.057)	(0.047)	(0.070)	(0.062)	(0.019)	(0.014)
Baseline Controls	-	+	-	+	-	+
N	27,584	27,167	15,053	14,880	7,793	7,643
b. Excluding Spain						
Presumed	-0.130**	-0.119**	-0.224***	-0.215***	-0.015	-0.012
	(0.057)	(0.047)	(0.071)	(0.063)	(0.019)	(0.014)
Baseline Controls	-	+	-	+	-	+
N	26,558	26,167	14,487	14,333	7,610	7,464

Standard errors (in parentheses) are clustered at country level, * $p < .1$, ** $p < .05$, *** $p < .01$

We observe significantly lower organ donation card holding in presumed consent countries on average. In line with our expectations, we find lower donation card holding among individuals who are willing to donate their organs in presumed consent countries. The negative coefficients stay statistically significant even after the control variables are accounted for. However, surprisingly among people who are not willing to donate, we do not observe higher registration in presumed consent countries. These findings imply that presumed consent legislation is likely to increase cadaveric donation rates because not willing individuals fail to register their preferences in presumed consent countries. This could be due to many reasons. For instance, informed consent countries might be more active in soliciting for organ donation through advertising donation card more eagerly. Another channel could be different registration costs in different countries. For instance, in the Netherlands[30] and in the UK[31] which are both informed consent countries, online registration of the

[30]From the following webpage: https://www.donorregister.nl/uwregistratie/campagne/
[31] From the following webpage: http://www.organdonation.nhs.uk/how_to_become_a_donor/

preference for organ donation is possible. On the other hand, Sweden[32] which is a presumed consent country also allows for online registration of the preference. But, difference in the costs of registering the preferences for organ donation between informed and presumed consent countries is beyond the scope of this study. For the purposes of this study, it suffices to observe higher organ donation card holding for both individuals who are willing to donate and not willing to donate in informed consent countries.

3.3.3 The Impact of Presumed Consent on Cadaveric Donors and Transplantations

3.3.3.1 Descriptive Statistics

Interestingly, Figure 3-1 shows that Bulgaria having the lowest cadaveric donor rate and Spain having the highest cadaveric donor rate are both presumed consent countries. This suggests that there are other factors behind relative efficiency of Spain and relative inefficiency in Bulgaria which has to be accounted.

Mainly following previous literature, information on a number of factors which could potentially impact organ donation rates such as number of deaths by specific causes, health spending, medical infrastructure, trust levels, religious beliefs and education level are gathered from a variety of sources[33]. In most cases, deceased donors were brain-dead and their hearts were artificially functioning with the help of ventilation machines. The most common causes of brain-death are homicides, motor vehicle accidents and cerebro-vascular diseases. We collected information for these causes of deaths from EUROSTAT. Information on health expenditure per capita and hospital beds per million populations as a proxy for transplant infrastructure are collected from WHO and EUROSTAT respectively. We also added the percentage of the population having upper secondary or tertiary education for each sex compiled from EUROSTAT.

According to Eurobarometer, 2009 survey on organ donation, religious reasons, distrust in the system and scare of manipulation of the human body are three major causes of refusal for organ donation. To capture trust in the system, corruption perceptions scores from Transparency International are included. Corruption perception scores ranges between 10 (highly clean) and 0

[32] From the following webpage: http://www.livsviktigt.se/Sa-har-tar-du-stallning/Sidor/default.aspx
[33] Detailed information about sources of the data is presented in Appendix B.

(highly corrupt). To control for religiosity changes over time, I compiled percentage of population being Roman Catholic and having no religion[34] mainly from four surveys. These surveys are International Social Survey Program (ISSP), European Social Survey (ESS), European Values Survey (EVS) and Eurobarometers conducted between 1999/2000 and 2010.

Table 3-6 provides summary statistics for the sample.

Table 3-6: Descriptive Statistics (means and standard deviations for 2000-2010)

	Entire Sample (1)	Presumed consent countries (2)	Informed consent countries (3)	Difference (S.E.) ((2)-(3))
Presumed consent country	0.74			
	[0.44]			
Cadaveric donor rate, pmp	15.96	16.95	13.17	3.78***
	[7.65]	[7.96]	[5.88]	(1.00)
Cadeveric donation after brain death, pmp	15.40	16.58	12.09	4.49***
	[7.36]	[7.65]	[5.24]	(0.95)
Kidney transplant rate, pmp	25.03	26.54	20.68	5.85***
	[11.38]	[11.75]	[8.94]	(1.48)
Homicide, pmp	21.48	21.64	21.01	0.63
	[25.80]	[25.96]	[25.48]	(3.45)
Deaths from motor vehicle accident rate, pmp	113.32	119.08	97.47	21.61***
	[48.99]	[42.60]	[60.91]	(6.45)
Deaths from cerebro-vascular diseases, pmp	853.95	867.44	816.88	50.55
	[498.73]	[473.35]	[564.44]	(66.86)
Health expenditures per capita	2238.78	2144.57	2507.97	-363.40*
	[1260.52]	[1236.08]	[1298.67]	(165.84)
Hospital beds, per 100,000 population	580.83	582.62	575.72	6.90
	[166.77]	[160.31]	[185.03]	(22.26)
Corruption perception score	6.28	5.98	7.12	-1.14***
	[1.99]	[1.84]	[2.16]	(0.26)
% of people having no religion	0.23	0.23	0.22	0.01
	[0.18]	[0.18]	[0.17]	(0.02)
% of people being roman catholic	0.42	0.45	0.34	0.12*
	[0.35]	[0.35]	[0.34]	(0.05)
% of higher educational attainment, males	68.62	67.45	71.95	-4.50**
	[13.13]	[14.62]	[6.39]	(1.72)
% of higher educational attainment, females	67.49	66.70	69.73	-3.03
	[12.07]	[13.38]	[6.74]	(1.59)
Number of Countries	27	20	7	

Notes: standard deviations in [ ]; standard errors in ( ).* $p < 0.05$,** $p < 0.01$,*** $p < 0.001$

The first column gives statistics for the entire sample while the second and third columns restrict the sample to presumed and informed consent countries. The fourth column provides information about the difference between the second and third columns and statistical significance of the difference. Presumed consent countries form 74 percent of the observations in our sample. The

[34] Includes people who consider themselves as agnostic, atheist or having no religion.

presumed consent countries have on average 16.95 cadaveric donors pmp per year, whereas informed consent countries have on average 13.17 cadaveric donor pmp per year. The fourth column shows that presumed consent countries have 3.8 more cadaveric donors per million populations per year and this difference is statistically significant. Cadaveric donation after brain death is even lower in informed consent countries. Similarly, presumed consent countries have statistically higher kidney transplant rate.

There are other differences between these set of countries. In particular, deaths from homicide, motor vehicle accidents and cerebro-vascular diseases are higher in presumed consent countries. However, only motor vehicle accident difference is statistically significant. Although health expenditure per capita are lower in presumed consent countries in comparison to informed consent countries, hospital beds per capita are higher. Yet, the difference in hospital beds is not statistically significant. Corruption perception scores of informed consent countries are higher than that of presumed consent countries. Since higher corruption perception scores means more transparency, we can say that informed consent countries are perceived as more transparent. The difference in corruption perception scores is statistically significant at 5% level. No major difference is observed in percentage of population considering themselves as having no religion while we observe higher percentage of population considering themselves as Roman Catholic in presumed consent countries. Relative to informed consent countries, presumed consent countries have statistically lower percentage of upper secondary or tertiary educated male population. On the other hand, the no significant difference was found for the upper-secondary or tertiary educational attainment of females.

3.3.3.2 Regression Output for Cadaveric Donation Rate

In this section, we provide regression analysis for total cadaveric donation rate and donation after circulatory death. We differentiate between donation after brain death (known as heart-beating donation) and donation after circulatory death (known as non-heart-beating donation) since donation after circulatory death less efficient for transplant outcomes other than kidneys. (Cota, Burgess, & English, 2013) Moreover, donors after circulatory death provide, on average, one fewer organ for

transplantation than donors after brain death. (NHS Blood and Transplant, 2012). Also, as a response to long waiting lists, more and more organs from circulatory death donors are used.

Table 3-7 shows regression output for log cadaveric donor rate (sum of donation after circulatory death and donation after brain death).

Table 3-7: Pooled OLS Estimates of Log Cadaveric Donor Rate

	(1)	(2)	(3)	(4)	(5)
Legislation					
Presumed consent	0.355	0.329**	0.390**	0.350*	0.042
	(0.401)	(0.157)	(0.148)	(0.190)	(0.163)
Practicing Legislation					
Family Consent					-0.309
					(0.341)
Presumed consent*Family Consent					0.552*
					(0.291)
Potential Donors					
Log of MVA+CVD+Homicide, pmp		-0.038	-0.072	-0.108	-0.072
		(0.076)	(0.076)	(0.072)	(0.068)
Health Spending					
Log of Health Expenditures per capita		0.227	0.212	0.190	0.190
		(0.181)	(0.185)	(0.212)	(0.162)
Medical Infrastructure					
Log of hospital beds per 100,000 people		-0.073	-0.226	-0.092	-0.151
		(0.273)	(0.258)	(0.280)	(0.241)
Trust in the system					
Corruption Perception Score		0.117***	0.160***	0.079	0.071
		(0.042)	(0.048)	(0.059)	(0.054)
Religious Beliefs					
% of people having no religion		1.824***	1.818***	1.914***	1.873***
		(0.345)	(0.347)	(0.508)	(0.487)
% of people being Roman Catholic		1.237***	1.278***	1.353***	1.381***
		(0.254)	(0.255)	(0.355)	(0.306)
Education					
% of higher educational attainment, males		-0.038**	-0.040**	-0.007	-0.000
		(0.017)	(0.015)	(0.020)	(0.019)
% of higher educational attainment, females		0.043**	0.045**	0.007	0.001
		(0.018)	(0.017)	(0.025)	(0.022)
Average willingness to donate			-1.153		
			(0.809)		
Country Group Fixed Effects	-	-	-	+	+
R-squared	0.037	0.655	0.664	0.741	0.755
N	287	262	262	262	262

Standard errors (in parentheses) are clustered at country level, * $p < .1$, ** $p < .05$, *** $p < .01$

Model (1) contains only a dummy variable for presumed consent legislation. According to model (1) cadaveric donor rate is 36% higher in presumed consent countries. However, the coefficient is not statistically significant. In model (2), to treat potential confounding factors, we include baseline control variables. These controls are death rates from homicides, motor vehicle accidents and cerebro-

vascular diseases, health expenditure per capita, hospital beds per 100,000 population, corruption perception scores, percentage of population considering themselves as having no religion, percentage of population considering themselves as Roman Catholic, percentage of upper secondary or tertiary educated males, and percentage of upper secondary or tertiary educated females. According to model (2), the coefficient of presumed consent variable indicates 33% higher cadaveric donor rate. This difference is significant at 5% level.

Model (3) includes average willingness to donate organs in each country from Eurobarometer surveys conducted in 2002, 2006 and 2009 to treat potential confounding effect of social acceptance of organ donation. According to model (3), the coefficient of presumed consent dummy variable is still significant after inclusion of average willingness for organ donation suggesting that the impact of presumed consent on cadaveric donor rate cannot be attributed to presumed consent being enacted in countries with higher social acceptance of organ donation.

Since there is no country which changed legislation in the period that we consider, we could not estimate country fixed effects which would treat any unobserved country level heterogeneity such as cultural inclination towards organ donation. To capture the idea that presumed consent might be enacted in countries where there is different unobserved heterogeneity, we group countries in model (4) according to following classification in which presumed consent countries are bolded.

(1) Ireland-UK,
(2) **Greece-Cyprus**,
(3) Germany -**Austria-Hungary** (Hungary is included in here because of the historical connection between Hungary and Austria.
(4) Netherlands-**Belgium-Luxembourg**
(5) **Estonia**-Lithuania-**Latvia** (These countries have formed Balttransplant organization among each other)
(6) Denmark-**Sweden-Finland** (They have formed Scandiatransplant organization among each other)
(7) **Poland-Czech Republic-Slovakia** (Western Slavic ethnic origin)
(8) **Portugal-Spain-Italy-France** (Latin ethnic origin)
(9) **Bulgaria**-Romania-**Croatia-Slovenia** (Southern-Slavic ethnic origin and also these countries are known as Balkan countries)

Model (4) also shows 35% higher donation rates in presumed consent countries when country group dummy variables are included in the model. These regression results suggest that presumed consent countries have 33-39% higher donation rates than informed consent countries on average. As

shown in the last row, model (4) explains almost three-fourth of the variance of the dependent variable (R^2 =0.741).

In model (5), we included family consent and its interaction with presumed consent to check whether family consent makes any difference since most of the presumed consent countries routinely seek family consent. Model (5) suggests that presumed consent does especially matter when it is combined with family consent, whereas in informed consent countries, seeking family consent has a negative but insignificant coefficient. This result is in line with the prediction that even if families make the final decision on organ donation, presumed consent laws may result in notably higher donation rates. This model suggests that presumed consent is primarily impacting cadaveric donor rate through reducing family refusals. Moreover, there is no marked differences in cadaveric donor rate between hard opt-out and soft-opt out countries.

The same five models of the cadaveric donor rate are also estimated for log of brain death donors rate. Table 3-8 provides regression output. Overall, the regression results suggest more pronounced impact of presumed consent.

Table 3-8: Pooled OLS Estimates of Log of Brain Death Donors Rate

	(1)	(2)	(3)	(4)	(5)
Legislation					
Presumed consent	0.408	0.386**	0.422***	0.366**	0.005
	(0.388)	(0.151)	(0.144)	(0.161)	(0.136)
Practicing Legislation					
Family Consent					-0.372
					(0.279)
Presumed consent*Family Consent					0.647**
					(0.245)
Baseline Controls	-	+	+	+	+
Country Group Fixed Effects	-	-	-	+	+
R-squared	0.050	0.665	0.668	0.762	0.781
N	287	262	262	262	262

Standard errors (in parentheses) are clustered at country level, * $p < .1$, ** $p < .05$, *** $p < .01$

3.3.3.3 Regression Output for Kidney Transplantation Rate

In this section, we provide regression analysis for the kidney transplantation rate. We used the same five models of the cadaveric donor rate for analyzing log kidney transplant rate. Table 3-9 provides regression output. Model (1) shows that presumed consent countries have 33% higher kidney transplant rate, however the coefficient is not statistically significant. Similar to cadaveric donor

regression results, model (2) and (3) reports that presumed consent countries have statistically significantly higher kidney transplant rate. In model (4), when country group dummy variables are included, presumed consent countries are found to have 31% higher kidney transplant rate. A similar pattern emerges from model (5) as in cadaveric donor rate. Presumed consent increases kidney transplant rate when it is combined with family consent. In line with intuition, family consent in informed countries has a negative; yet insignificant coefficient.

Table 3-9: Pooled OLS Estimates of Log Kidney Transplant Rates

	(1)	(2)	(3)	(4)	(5)
Legislation					
Presumed consent	0.328	0.347**	0.367**	0.309*	0.015
	(0.387)	(0.154)	(0.153)	(0.178)	(0.141)
Practicing Legislation					
Family Consent					-0.180
					(0.315)
Presumed consent*Family Consent					0.525*
					(0.278)
Baseline Controls	-	+	+	+	+
Country Group Fixed Effects	-	-	-	+	+
R-squared	0.030	0.614	0.615	0.737	0.756
N	296	270	270	270	270

Standard errors (in parentheses) are clustered at country level, * $p < .1$, ** $p < .05$, *** $p < .01$

Overall, the regression results from Table 3-7 and 3-9 suggest that presumed consent countries have 28 to 32% higher cadaveric donation and 27 to 31% higher kidney transplant rates in comparison to informed consent countries.

We conducted a number of robustness checks such as inclusion of year fixed effects and exclusion of Spain from the sample. Weighted least squares regression is also estimated in which weighting proportional to size of population is used together with adjustments according to absolute value of the residual. Details of these tests are reported in Appendix C.

3.4 Discussion

In this section, we discuss whether the observed effects were attributable to presumed consent effects. For instance, Healy (2005) suggests that presumed consent is an indicator of a country's commitment to organ donation rather than a causal mechanism in itself. Secondly, we discuss potential problems with both types of consent legislation.

A country's commitment to organ donation could be indicated by available transplant capacity. In our main analysis, transplantation capacity is imperfectly proxied by number of hospital beds per 100,000 people. We include natural logarithm of number of kidney transplant centers per million populations as an additional control variable in Table 3-10. Model (1) and (2) in Table 3-10 uses the same specifications as of model (3) and (4) in Table 3-7 with logarithm of number of transplantation centers per million population as an additional control variable. Table 3-10 reports increased coefficient of presumed consent compared to Table 3-7. This finding suggests that commitment to organ transplantation proxied by kidney transplant infrastructure is negatively correlated with presumed consent legislation.

Kidney transplant capacity (defined as the number of kidney transplant centers per million populations) could also influence kidney transplantation rates. Therefore; we also examine how inclusion of kidney transplant centers changes the kidney transplantation rate regression results. Model (3) and (4) in Table 3-10 uses the same specifications as of model (3) and (4) in Table 3-9 with logarithm of number of transplantation centers per million population as an additional control variable. Models (3) and (4) also reports increased coefficient of presumed consent compared to Table 3-9. This finding also suggests negative correlation between transplant commitment proxied by kidney transplant capacity and presumed consent legislation.

Table 3-10: Transplant Infrastructure and Presumed Consent (Pooled OLS)

	Log Cadaveric Donor Rate		Log Kidney Transplant Rate	
	(1)	(2)	(3)	(4)
Legislation				
Presumed consent	0.546***	0.683***	0.494***	0.561***
	(0.147)	(0.164)	(0.139)	(0.161)
Transplant Infrastructure				
Log of kidney transplant centers	0.299**	0.274	0.352**	0.322**
	(0.128)	(0.172)	(0.141)	(0.153)
Baseline Controls	+	+	+	+
Average willingness to donate	+	-	+	-
Country Group Fixed Effects	-	+	-	+
R-squared	0.778	0.828	0.772	0.836
N	153	153	153	153

Standard errors (in parentheses) are clustered at country level, * $p < .1$, ** $p < .05$, *** $p < .01$

On the whole, the results of Table 3-10 suggest that the regression results provided in Table 3-7 and 3-9 are not confounded by commitment to organ donation differences across countries.

Concerning disadvantages, both presumed consent and informed consent legislation are error-prone. In presumed consent legislation, non-willing individuals' organs could be mistakenly removed if they did not register their non-willingness. (Gill, 2004; Orentlicher, 2008) In this case, removal of an organ is wrong because it is against the patient's will. In informed consent countries, willing individuals' organs could be mistakenly not removed if they did not register their willingness. (Gill, 2004; Orentlicher, 2008) In this case, non-removal of an organ is a waste of a scarce resource which could be used for improving life of another person. So, in both types, there is a possibility that some people's wishes are not respected.

Using descriptive information from Table 3-2 and Table 3-4, we can calculate which consent regime produces the least percentage of errors. Similarly, we can calculate which consent regime maximizes the percentage of people whose wishes are respected. As discussed, socially desirable responding would not change our conclusions as long as individuals from informed and presumed consent regimes are not different regard to their engagement in socially desirable responding.

The probability of making mistake in presumed consent legislation is 43% which is calculated by multiplying the percentage of non-willing individuals with the probability of non-registering their preference (= (100-55.83)*(100-1.56)/100). The probability of making mistake in informed consent legislation is 37.71 which is given by multiplying the percentage of willing individuals with the probability of non-registering their preference (=55.56*(100-32.11)/100). If the default consent regime should be chosen such that the least number of errors are made, informed consent regime seems to be superior.

The percentage of people whose wishes are respected in informed consent regime is the ones who are not willing to donate (25.45%) plus the ones who are willing to donate and registered their preference (=55.56*32.11/100). This corresponds to 43.29%. The percentage of people whose wishes are respected in presumed consent regime is the ones who are willing to donate (55.83%) plus the ones who are not willing to donate and registered their preference (=26.99*1.56/100). This corresponds to 56.25%. If the criterion for choosing the default consent regime is to maximize the percentage of people whose wishes are respected, presumed consent seems to outperform inform consent regime.

3.5 Conclusion

How to produce larger number of organ donors is a relevant question since the chronic shortage of human organs is leading to loss of many patients while waiting for an organ in Europe as in other places. One proposition by medical professionals and some politicians in informed consent countries such as the UK and Netherlands is changing legislative defaults on organ donation to presumed consent. However, less is known whether presumed consent is effective for increasing cadaveric donation rates. Some previous studies find higher organ donation rates in presumed consent countries, the evidence was not quite convincing. Moreover, there is no consensus in the literature on whether the presumed consent legislation is an indicator of a country's commitment to organ donation or a causal mechanism in itself.

In this study, we attempt to extend the literature on how presumed consent impacts cadaveric donors and kidney transplantations in the EU-27 countries plus Croatia in the period 2000-2010. As a first step, we show evidence that presumed consent is not necessarily legislated in countries where there is higher social acceptance of organ donation. For presumed consent to be effective, individuals should fail to register their preferences in case of a mismatch with their preference and legislative default. In the second step, we show that people fail to take action in line with their preferences, this is especially more relevant for individuals who are non-willing to donate their organs. Therefore, we suggest that presumed consent is likely to produce higher organ donors. In the main analysis, after accounting for potential confounding factors, our estimates suggest that presumed consent countries have 28 to 32% higher cadaveric donation and 27 to 31% higher kidney transplant rates in comparison to informed consent countries.

Changing defaults to presumed consent has its advantages and disadvantages. Although our study indicates that presumed consent legislation can be instrumental for saving lives, we do not claim that changing the system would be ideal from every aspect. If policy makers attribute more importance to minimizing the number of errors, informed consent seems to be doing better.

Appendix A: Legislation and Practices on Cadaveric Organ Donation

Country	Presumed Consent	Source	Routine Family Consent	Source
Austria	Yes	GODT[35]	No	GODT
Belgium	Yes	(Gevers, Janssen, & Friele, 2004)	Yes	(Gevers, Janssen, & Friele, 2004)
Bulgaria	Yes	GODT	Yes	GODT
Croatia	Yes	GODT	Yes	GODT
Cyprus	Yes	GODT	Yes	GODT
Czech Republic	Yes	GODT	No	GODT
Denmark	No	(Abadie & Gay, 2005)	No	(Bilgel, 2010)
Estonia	Yes	GODT	Yes	GODT
Finland	Yes	(Abadie & Gay, 2005)	Yes	GODT
France	Yes	GODT	Yes	GODT
Germany	No	GODT	No	GODT
Greece	Yes	(Abadie & Gay, 2005)	Yes	GODT
Hungary	Yes	GODT	Yes	GODT
Ireland	No	GODT	Yes	GODT
Italy	Yes	GODT	Yes	GODT
Latvia	Yes	GODT	No	(Paparde, 2010)
Lithuania	No	GODT	Yes	GODT
Luxembourg	Yes	GODT	No	GODT
Netherlands	No	(Abadie & Gay, 2005)	Yes	(Bilgel, 2010)
Poland	Yes	(Abadie & Gay, 2005)	No	(Bilgel, 2010)
Portugal	Yes	GODT	Yes	GODT
Romania	No	GODT	Yes	GODT
Slovakia	Yes	GODT	Yes	GODT
Slovenia	Yes	(Abadie & Gay, 2005)	Yes	GODT
Spain	Yes	GODT	Yes	GODT
Sweden	Yes	(Abadie & Gay, 2005)	No	GODT
United Kingdom	No	GODT	Yes	GODT

[35] (Global Observatory on Donation and Transplantation, 2011)

Appendix B: Sources of Data

B.1. Aggregate Country Level Data

The data for cadaveric organ donation rates are mainly from The International Registry in Organ Donation and Transplantation (IRODaT) of Transplantation Procurement Management. If the value is is missing in IRODaT, we checked Transplant Newsletters of corresponding years and used the corresponding values. The data for kidney transplantations are also from The International Registry in Organ Donation and Transplantation (IRODaT) of Transplantation Procurement Management. If the value is is missing in IRODaT, we checked Global Transplant Observatory and used the corresponding values. If the value is also not available in Global Transplant Observatory, we also checked Transplant Newsletters of corresponding years and used the corresponding values.

Homicide rates are available only till 2010 from EUROSTAT webpage. The missing figures from EUROSTAT are replaced by available figures from UN's homicide database. Deaths caused by Cebrebro vascular diseases for all ages and all sexes are available untill 2010 from EUROSTAT webpage. The missing values from EUROSTAT are replaced by available values from WHO's World Mortality Database. Deaths caused by transport accidents rates are also obtained from EUROSTAT's webpage. The missing values from EUROSTAT are replaced by available values from WHO's Mortality Database.

For health spending, total expenditure on health per capita PPP is obtained from the WHO's webpage. Hospital beds per 1000 people are obtained from EUROSTAT's webpage. The missing values are replaced by available values from the World Bank Database's webpage.

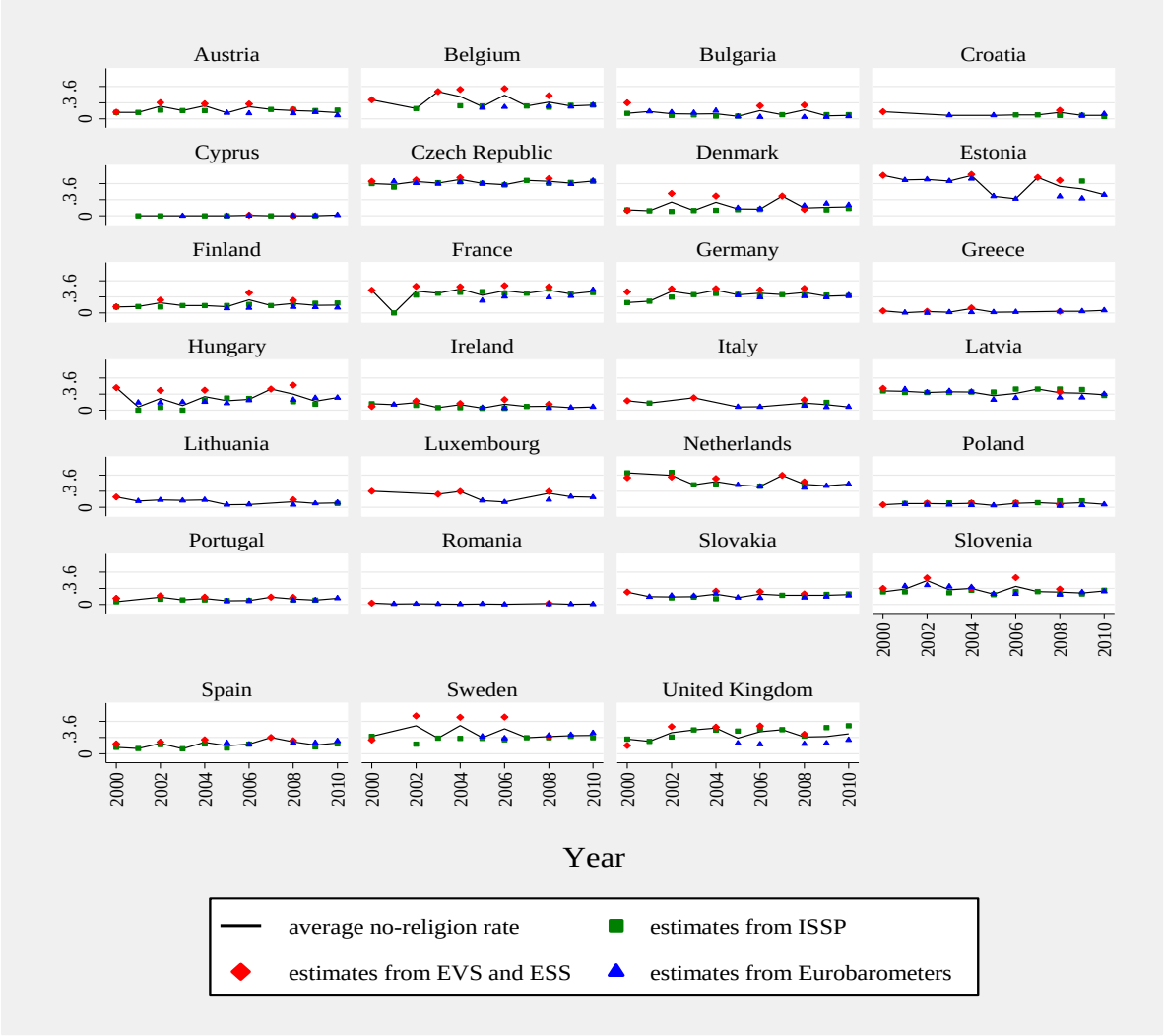
For trust in the system, corruption perception scores compiled by Transparency International are obtained from the webpage. For education, persons with upper secondary or tertiary educational attainments for both sexes are obtained from EUROSTAT's webpage.

For capturing religious preference changes over time, percentage of population being Roman Catholic and having no religion are compiled from mainly International Social Survey Program (ISSP), European Social Survey (ESS), European Values Survey (EVS) and Eurobarometers conducted between 1999/2000 and 2010. People having no religion are the people who consider themselves as agnostic, atheist or having no religion. If more than one value is obtained for the same

country for the same year from different surveys, weighted average of these values is used. Figure 3-3 shows no religion rate estimates for each country in the sample from the surveys. In general, Figure 3-3 shows small shifts or no shifts over time probably due the short time span covered. Except for six countries (Belgium, Estonia, France, the Netherlands, Sweden and United Kingdom), estimates from three surveys show a consistent pattern.

The number of kidney transplant centers is obtained from Transplant Newsletters of Spanish Organ Transplantations Agency. The information was available from 2003 to 2011.

Figure 3-3: Percentage of Considering Themselves as Having No Religion over Time by Country



B.2. Eurobarometer Surveys

Firstly, we use individual level data from Eurobarometer Surveys conducted in 2002, 2006 and 2009. The first one is Eurobarometer 58.2: Health and Developing Countries. This survey has respondents from 15 EU countries before the enlargement in 2004. The sample size is 16,230 individuals. The second one is Eurobarometer 66.2: Nuclear Energy and Safety, and Public Health Issues. This survey has respondents from all 27 EU countries also from Croatia and the Turkish Cypriot Community. The sample size is 28,584 individuals. The last one is Eurobarometer 72.3: Public Health Attitudes, Behavior, and Prevention. This survey has respondents from all 27 EU countries, Croatia, Turkey, Turkish Cypriot Community and Macedonia (FYROM). The sample size is 30,292 individuals.

The Eurobarometer surveys is based on multi-stage national probability samples of the citizens of participating countries aged 15 and over and carried out on behalf of the European Commission. The surveys are representative of the whole territory of the countries included in the analysis. These surveys are available from Leibniz Institute for Social Sciences.

Appendix C: Robustness Checks

Table C-1: Pooled OLS Estimates of Log Cadeveric Donor Rate (Excluding Spain)

	(1)	(2)	(3)	(4)	(5)
Presumed consent	0.306	0.304*	0.355**	0.344*	0.057
	(0.401)	(0.154)	(0.146)	(0.193)	(0.155)
Family Consent					-0.267
					(0.339)
Presumed consent*Family Consent					0.515*
					(0.284)
Baseline Controls	-	+	+	+	+
Country Group Fixed Effects	-	-	-	+	+
R-squared	0.029	0.645	0.651	0.735	0.749
N	276	251	251	251	251

Standard errors (in parentheses) are clustered at country level, * $p < .1$, ** $p < .05$, *** $p < .01$

Table C-2: Pooled OLS Estimates of Log Cadeveric Donor Rate (Year Fixed Effects)

	(1)	(2)	(3)	(4)	(5)
Presumed consent	0.355	0.325**	0.394**	0.340*	0.003
	(0.401)	(0.158)	(0.146)	(0.189)	(0.175)
Family Consent					-0.391
					(0.381)
Presumed consent*Family Consent					0.603*
					(0.325)
Baseline Controls	-	+	+	+	+
Country Group Fixed Effects	-	-	-	+	+
R-squared	0.037	0.666	0.678	0.749	0.763
N	287	262	262	262	262

Standard errors (in parentheses) are clustered at country level, * $p < .1$, ** $p < .05$, *** $p < .01$

Table C-3: Weighted OLS Estimates of Log Cadaveric Donor Rate

	(1)	(2)	(3)	(4)	(5)
Presumed consent	0.356***	0.272***	0.351***	0.378***	0.075
	(0.107)	(0.068)	(0.073)	(0.075)	(0.112)
Family Consent					-0.289*
					(0.157)
Presumed consent*Family Consent					0.565***
					(0.165)
Baseline Controls	-	+	+	+	+
Country Group Fixed Effects	-	-	-	+	+
R-squared	0.038	0.675	0.687	0.763	0.778
N	287	262	262	262	262

Table C-4: Pooled OLS Estimates of Log Kidney Transplant Rate (Excluding Spain)

	(1)	(2)	(3)	(4)	(5)
Presumed consent	0.289	0.329**	0.340**	0.306	0.025
	(0.388)	(0.152)	(0.154)	(0.180)	(0.137)
Family Consent					-0.152
					(0.314)
Presumed consent*Family Consent					0.501*
					(0.274)
Baseline Controls	-	+	+	+	+
Country Group Fixed Effects	-	-	-	+	+
R-squared	0.024	0.605	0.605	0.731	0.751
N	285	259	259	259	259

Standard errors (in parentheses) are clustered at country level, * $p < .1$, ** $p < .05$, *** $p < .01$

Table C-5: Pooled OLS Estimates of Log Kidney Transplant Rate (Year Fixed Effects)

	(1)	(2)	(3)	(4)	(5)
Presumed consent	0.328	0.341**	0.368**	0.300	-0.062
	(0.387)	(0.158)	(0.153)	(0.182)	(0.140)
Family Consent					-0.347
					(0.379)
Presumed consent*Family Consent					0.646**
					(0.306)
Baseline Controls	-	+	+	+	+
Country Group Fixed Effects	-	-	-	+	+
R-squared	0.030	0.619	0.621	0.745	0.764
N	296	270	270	270	270

Standard errors (in parentheses) are clustered at country level, * $p < .1$, ** $p < .05$, *** $p < .01$

Table C-6: Weighted OLS Estimates of Log Kidney Transplant Rate

	(1)	(2)	(3)	(4)	(5)
Presumed consent	0.325***	0.288***	0.316***	0.370***	0.081
	(0.103)	(0.067)	(0.076)	(0.075)	(0.112)
Family Consent					-0.134
					(0.157)
Presumed consent*Family Consent					0.501***
					(0.167)
Baseline Controls	-	+	+	+	+
Country Group Fixed Effects	-	-	-	+	+
R-squared	0.030	0.619	0.621	0.745	0.764
N	296	270	270	270	270

4 Donate More, Be Happier!

4.1 Introduction

Making donations is a form of pro-social behavior. Why people donate is an interesting question from an economics point of view since people give away their money for free. The economics literature considers a number of motivations, including signaling one's social status, enhancing one's reputation, and acquiring tax-break advantages (Glazer & Kondrad 1996; Harbaugh 1998; Clotfelter 1985). Some argue that people have a taste for giving; they experience a "warm glow" from having "done their bit" (Andreoni, 1989; 1990).

To put the warm glow motivation into context, we have to make a distinction between pure altruism and impure altruism at this point. Economists describe a person as an altruist if other peoples' welfare enters into that person's utility function. A person is pure altruist if she only cares for the final situation of the other person regardless of what she personally did for the other person, whereas an impure altruist would enjoy not only the final situation but also her own deed.

In large economies, under the assumption of non-cooperative equilibria, warm-glow motive must dominate at the margin (Ribar & Wilhelm, 2002). The intuition is that the incentive to free ride must be so overwhelming if large numbers of others are collectively providing a substantial amount of charity, the only justification for giving is that donors get some direct benefit from giving.

There are several channels why humans may have pleasurable psychological experiences by acting pro-socially. Firstly, humans feel sympathy or empathy for the needs of another human being which might drive pro-social behavior. Being aware of those in need and ignoring any possible help would result in feelings of shame and guilt. Pro-social behavior may alleviate those feelings. Moreover, pro-social behavior can make one feel proud of oneself by acting in line with a certain self-image or social norm.

Although the "warm glow" hypothesis has existed since the 1980s, to the author's knowledge, in the economics literature, no study measures the magnitude of the effect of donating on subjective wellbeing using real donation amounts. Previous economics literature on "warm glow" is largely theoretical; existing empirical work focuses on measuring crowding out to demonstrate the existence

of impure altruism or warm glow (Andreoni 1989; 1990; Crumpler and Grossman 2008). Therefore, whether donating makes people happier in a causal way remains unanswered, partly because happiness is not thought to be in the area of economics. However, Stiglitz, Sen and Fitoussi (2009) argue that it is possible to collect meaningful and reliable data on both subjective and objective well-being. Subjective wellbeing encompasses evaluation of one's life and happiness which are collected by questions on how one feels in surveys.

Psychologists have worked on the relationship between happiness and pro-social behavior and have produced three strands of literature. One argues that pro-social behavior promotes happiness. Donors self-report "feeling good" as a motive for donating to charitable causes. In Wunderink (2002), 57% of respondents mention "gives me a good feeling" as a motivation for making donations in the Netherlands. Another strand produces experimental evidence that happiness increases charitable behavior. Researchers find participants more likely to help others after experiencing a positive mood (Aderman, 1972; Rosenhan, Underwood and Moore, 1974; Forgas, Dunn and Granland, 2008). Finally, some researchers say that certain personality characteristics might foster both giving and happiness. For instance, in Bekkers (2006), extroverted individuals who are generally happier are also found to be more likely to give.

The psychology literature exploring the benefits of charitable giving is largely experimental and has conducted on small groups of students or volunteers, including Field et al. (1998), Lyubomirsky, Sheldon and Schkade (2005) and Dunn, Aknin and Norton (2008), leaving open the question of the external validity of results.

Understanding donating is important as donations matter economically. The total value of monetary and goods donated by Dutch households in 2009 amounts to € 1,938 million (Schuyt, Gouwenberg, & Bekkers, 2011). The average donation by all households in 2009 in the form of cash and goods is calculated to be €210. This issue is relevant, since 87% of Dutch households donate money to charitable organizations[36] (Schuyt, Gouwenberg, & Bekkers, 2011).

[36] In 2010, 41% of the Dutch population was engaged as a volunteer in unpaid work for a social organization at least once in the previous year (Schuyt, Gouwenberg, & Bekkers, 2011).

In this study, we examine whether charitable giving makes people happier taking into account the potential reverse causality and personality related fixed factors. For this purpose, we employ two panel datasets from the Netherlands: five waves of the Longitudinal Internet Studies for the Social Sciences (LISS) panel and four waves of the Giving in the Netherlands Panel (GINP). The LISS includes a happiness score for each wave from 2007 to 2012. In the GINP, a life satisfaction measure is available for the 2006 wave; self-reported health is available in four waves measured biannually from 2004 to 2010. To avoid dropping observations, we use self-reported health as a proxy for subjective wellbeing. The GINP asks whether the respondent has made a donation in the previous two weeks; the LISS panel asks about donations in the last 12 months. This difference makes the GINP more useful to detect short-term effects.

Exploiting the panel nature of both the LISS panel and the GINP, we initially run fixed effects models. However, fixed effects specification is not enough to say something about direction of causality. To identify the direction of causality, we employ an instrumental variables approach. We use different types of personal solicitation as instruments for donating since the literature suggests that solicitation is an effective way to induce people to make charitable donations (Meer and Rosen 2011; Bekkers 2005; Van Diepen, Donkers and Franses 2006).

The GINP panel asks respondents whether they have been solicited in the last two weeks by 13 different fundraising methods. If we think about donation as a normal good, donors constitute the supply side of this market and charities constitute the demand side of the market. Using "solicitation" as instrument is plausible since it comes from the demand side of the market. Also, the solicitation of money puts social pressure on prospective donors and increases the likelihood that they will agree to donate. Although a prospective donor may want to donate, she or he may not because of an information gap. For instance, they might not know who needs money the most and how to send it. Charities through solicitation can supply this information. Also, successful solicitation strategies may enhance the probability of giving by raising awareness of needs for donation and ensuring donors that their gift makes a difference.

The 13 types of solicitation as our instrument set are jointly significant in explaining donation behavior as the first stage results show. However, the validity of our instrumental variable (IV) results

rests on the assumption that types of solicitation does not affect subjective wellbeing directly. The main concern with using personal solicitation information is that charities might target some people rather than others and this might be related to subjective wellbeing in an unobserved way. Charities are known to concentrate their solicitation efforts on affluent potential contributors. Since charities generally target people with higher economic wellbeing, this does not suggest that there is targeting on the basis of psychological wellbeing because evidence suggests that money does not necessarily lead to greater happiness after a certain threshold (Kahneman & Deaton, 2010). Also, I do control for indicators of income in the regression model.

From the fixed effect estimates of the LISS panel, which primarily measure long-term effects, we find evidence that engaging in donation is associated with higher happiness scores. From the fixed effect estimates of the GINP, which show short-run effects, we do not find any significant relationship between donating and subjective health. We find a concave relationship between donating and life satisfaction using the OLS for the 2006 wave of GINP. For the IV estimates using GINP, contrary to intuition, we find negative effect of donating on subjective health. When experience of certain diseases is taken into account, the effect is no longer significant. This suggests that being solicited is negatively correlated with health status. In line with our expectation, we find a significant effect of donating on life satisfaction using an IV methodology for the 2006 wave GINP. Solicitation reflecting selection cannot explain the results that we observe for life-satisfaction.

To put the findings from IV estimation, we can compare the coefficient of income from employment and coefficient of amound donated. This comparison suggest that an increase of €1 in donations increases life satisfaction as much as a €104 increase in income from employment for those who was affected by solicitation. At first, the effect might seem implausible. However, this is the local average treatment effect (LATE) for individuals who donate an extra Euro because they are solicited and these individuals would not donate this extra Euro if they were not solicited. This group of individuals is not likely to be representative of the Dutch population. Thus, IV estimates might not reflect the average treatment effect. Moreover, the opt-in nature of the GINPS sampling also casts doubt on the effect size calculation. Therefore, extrapolation is not meaningful.

For discussing the policy implications, the average treatment effect would be more useful. Nevertheless, the main message of this study - donation makes at least some people happier-, could have policy implications for boosting charitable giving. Under rationality, people are thought to make optimal decisions in which they are supposed to take into account psychological benefit of donation in their utility function. At first, since a large number of people already donated and people can learn the psychological benefits of donation over time, rationality seems to be a valid assumption. Rationality does not leave room for advertising psychological benefits of donation to affect donation amounts. However, evidence reveals that people overlook the benefits of charitable giving. Dunn, Aknin and Norton (2008) showed that people erroneously thought that personal spending would make them happier than pro-social spending although they found higher happiness levels of randomly assigned pro-social spenders. Frank (2004) discusses the evidence on how people do not spend their money in ways that yield significant and lasting increases in measured satisfaction. Therefore, there might be still room for increasing subjective wellbeing by engaging in pro-social behavior.

4.2 Literature Review

As noted above, there are three strands of literature on the relationship between pro-social behavior[37] and subjective wellbeing: one argues that being pro-social promotes happiness; the second says happier people are more likely to be pro-social; the third argues that personality traits is driving happiness and pro-social behavior. [38].

The first strand dates back to Aristotle's concept of eudemonia which could be summarized as happiness from performing moral duties. Recent evidence from neuropsychological studies suggests that donations to charity "elicit neural activity similar to experience of pleasure" (Harbaugh, Myer and

[37] There are different forms of pro-social behavior. In this study, we focus on donating money.

[38] Dolan, Peasgood, & White (2008) provide a brief overview of the literature on volunteering and subjective well-being relationship. Two studies from this literature are Meier and Stutzer (2008), Field et al (1998). Meier and Stutzer (2008) find higher levels of volunteer work to be associated with higher levels of overall life satisfaction. They use German Socioeconomic Panel data after the fall of the Berlin Wall but before German reunification when volunteering opportunities dropped dramatically in East Germany. This quasi-experiment allows them to compare happiness of East Germans to a control group who did not experience a change in their volunteering status. Field et al. (1998) find that elderly volunteers experience less anxiety and depression and enjoy improved health after being asked to give infants a massage. Fully covering the literature on volunteering and subjective wellbeing is beyond the scope of this study and not directly relevant for measuring the warm glow effect of donating.

Burghart 2007; Tankersley, Stowe and Huettel 2007; Moll et al. 2006). Social psychologists have shown that acting pro-socially may contribute to less anxiety and depression and a more positive mood. Lyubomirsky, Sheldon and Schkade (2005) show that asking people to commit random acts of kindness can significantly increase happiness levels for several weeks. Spending money on others is also found to increase happiness more than spending money on oneself (Dunn, Aknin, & Norton, 2008). Researchers randomly assigned some students to spend money on others and some to spend money on themselves; the former were significantly happier at the end of the day. However, Dunn, Aknin, & Norton (2008)'s are not really capturing the subjective wellbeing effect of donations since donation by definition should be voluntary.

In the second strand, experimental evidence shows that happiness increases charitable behavior. The majority of the experimental research has included mood inductions such as having participants read mood inducing statements (Aderman, 1972), recalling mood appropriate memories (Rosenhan, Underwood, & Moore, 1974), receiving cookies (Isen & Levin, 1972). Researchers find participants more likely to help others after experiencing a positive mood. A positive mood also increases helpful behavior in the workplace (Forgas, Dunn, & Granland, 2008). Using survey data, Wang and Graddy (2008) find that happiness affects religious giving but not secular giving; they argue that happy people are more emotionally able to help others and have more optimistic personalities, thus fostering charitable giving. In the workplace, employees who report being in a good mood are more likely to display helpful behaviors that are not part of their formal job requirements (Williams & Shiaw, 1999).

A third strand argues that certain personality characteristics might foster both giving and happiness. A survey in the United Kingdom finds that people who report a stronger sense of accomplishment are more likely to donate (Sargeant, Ford, & West, 2000). A study in New Zealand finds that people with a passive life orientation are less likely to donate (Todd & Lawson, 1999). A similar finding emerges from a study in the Netherlands in which more extroverted individuals are more likely to give and to give higher amounts (Bekkers, 2006). There is also growing evidence in psychological research that personality traits such as extroversion are important predictors of reported well-being (De Neve & Cooper, 1999).

In short, evidence on the effect of donating on subjective wellbeing has been largely based on experimental studies with relatively small sample sizes and amounts donated; whether donating actually causes greater happiness remains partly unanswered.

While no study has used solicitation an instrument, many consider the stimulant effect of solicitation on donation. For instance, there is evidence that giving typically occurs in response to solicitation (Bekkers & Wiepking, 2007). One study finds 85% of gifts are made following solicitation (Bryant et al. 2003); similarly, Bekkers (2005) finds that 86% of donations follow solicitation. Yoruk (2009) documents a positive relationship between personal solicitation and total charitable giving. Van Diepen, Donkers and Franses (2006) find that direct mailing solicitations increase giving to a group of charitable institutions in the Netherlands. A similar finding is observed for alumni donations in natural experiment setting (Meer & Rosen, 2011). The evidence that solicitations enhance the likelihood of donating is complemented by field work showing that personal solicitations are more effective than anonymous fundraising methods (Landry et al. 2006; Alpizar, Carlson and Johansson-Stenman 2008).

4.3 Empirical Framework

We are interested in estimating the impact of pro-social behavior (i.e., donating) on happiness. Let Y represent an outcome such as a happiness score, subjective health score or life satisfaction score, X represent observable personal characteristics, and PS denote pro-social behavior. For person "i", then, the model can be written as

$$Y_{it} = X_{it}\beta + \alpha_i PS_{it} + \varepsilon_{it} \quad [1]$$

$$PS_{it} = X_{it}\,\gamma + Z_i\delta + \theta_{it} \quad [2]$$

The treatment effect α_i will be positive if pro-social behavior increases happiness after controlling for an extensive set of covariates denoted by X in equation (1).

Pro-social behavior will be affected by observable personal characteristics, X_{it}, unobserved characteristics, θ, and by the activities of charitable organizations, Z_i as in equation (2). As mentioned in Section 1, we use engaging in donation and the amount donated as indicators for pro-

social behavior. As donating means giving money away, this necessarily imposes financial costs on donors. Therefore, an individual's financial resources are potentially correlated with the amount of donation. To account for potential correlations, employment status and net monthly income from employment variables are included in X. We include indicators for volunteering in X to treat possible correlations between volunteers and donors; if generous people are more likely to donate, they might also be more likely to volunteer. One could argue that the coefficient of donating or volunteering might indicate something about sociability of the respondent rather than the effect of donating or volunteering *per se*. Therefore, being a member of a social organization and participating in an event in social organization are included in X as independent variables to account for sociability differences.

One critical issue in estimating the impact of pro-social behavior on subjective wellbeing is that subjective wellbeing levels might not be comparable between individuals. In this case, cross-sectional analysis is invalid. Another issue is self-selection of individuals into being pro-social. In other words, if certain personality types are more likely to be pro-social and also to report higher subjective wellbeing, then the difference in reported subjective wellbeing scores between donors and non-donors may be due to unobserved personality types. To address these issues, we would need panel data since with these, the same person's subjective wellbeing could be compared over time. Moreover, time-invariant personality factors are cancelled out by examining the same person over time. Thus, fixed effects models will give an unbiased estimate of the effect of donating on subjective health if the only source of endogeneity is fixed personality related heterogeneity.

As reverse causality is also an issue (there is evidence that induced happiness makes people more likely to donate), instrumental variable (IV) techniques should be used. The estimation of IV requires an instrument which gives an exogenous variation of pro-social behavior. That is, we look for instruments which will impact a person's decision to be pro-social without affecting his or her subjective wellbeing directly. For instance, solicitation could be an instrument for donation since some people are asked to donate to charitable causes while others are not. The solicitation instrument set that may potentially affect donation behavior contains fourteen binary variables. These are as follows: Person was asked to donate (1) via door to door collection; (2) via street collection; (3) via sponsor campaign; (4) via collection in the church; (5) via a collection at work; (6) via television

campaign or telethon; (7) via a direct mail letter; (8) via internet/e-mail; (9) via collection during an event; (10) via collection through membership organization; (11) via advertisement; (12) via buying something; (13) via lottery tickets.

For our IV estimates to be consistent and valid, three conditions must hold. First, the instruments should be "relevant" for donating decision. Secondly, exclusion restrictions must hold. Thirdly, monotonicity should not be violated.

For the first condition, there are several reasons for solicitation being a relevant determinant of donating. The solicitation of money puts social pressure on the individuals asked and increases the likelihood that they will agree to donate. Successful solicitation strategies may enhance the probability of giving by raising awareness of needs and assuring donors that their gift makes a difference. Some people donate even though they are not asked to do so, but we expect that the propensities to donate will be lower than among those who are asked.

We expect a positive relationship between the solicitation and the ensuing donation. However, that is not enough. We need to make sure that instruments are highly correlated with the variable they instrument for, in this case, amount donated and donated dummy. This is commonly judged by examining an F-test on the instruments in a regression of the endogenous variable on the instruments (the first stage). Bound, Jaeger and Baker (1995) suggested that this F-statistic should be large and statistically significant; as a rule of thumb, Staiger and Stock (1997) say that an F-statistic of less than 10 could signal weak instruments.

Secondly, exclusion restrictions must also hold; instruments should be uncorrelated with the unobserved characteristics that can affect subjective wellbeing. We need to make sure that solicitation does not impact happiness directly. People might not like to be asked for money, and this might affect their happiness negatively; if so, our results might be the lower bound of the effect of donation on subjective wellbeing. However, the exclusion restriction would also be violated in the case of targeting by charities. To address this concern, we need to know how charities work.

4.3.1 How do charities work?

Although charities are non-profit organizations, their strategies for fundraising are more like enterprises: they work with direct mailings, the internet, and television. To increase their funds, they frequently try to promote donating through solicitation. The fundraising process employs both impersonal methods such as television campaigns and personal methods such as collection via a membership organization. Both types of solicitation increase the likelihood of contributions. For our analysis to be valid, we must ask and answer the following question: "Does solicitation matter because it stimulates potential donors or because it targets those who are already likely to donate"

Solicitation as a selection strategy: Given limited time and resources, charities cannot possibly solicit everybody to donate for a particular goal. As a result, seeking to maximize both efficiency and the probability of donation, they target those with donation potential. To this end, many charitable organizations keep detailed records of all individuals who have donated to their charity in the past *(Jonker, Paap, & Franses, 2000)*. They draw heavily on their own mailing lists to improve target selection. Potential donors are often selected based on their behavior in the past. When direct mailing is used, often Recency, Frequency, and Monetary value (RFM) are the only criteria for target selection *(Jonker, Paap, & Franses, 2000)*. Recency measures duration since an individual's last donation. Frequency counts an individual's positive responses to mailings during a certain period of time. Lastly, monetary value shows the amount of donation during a certain period of time. The RFM variables are used in direct marketing techniques. They are often combined into an individual score, which is then used to rank the individuals who are most likely to respond *(Bauer, 1988)*. In other words, individuals targeted for solicitation tend to be more likely to donate in the first place. Those solicited are likely have a higher income, be better educated, and have wide networks in the community.

Solicitation as a Stimulus: Solicitation may also influence contribution decisions by acting as a stimulus that mobilizes potential contributors and increases the likelihood of contributing in ways that are independent of targeting. Solicitations for money reduce the information costs associated with contributing by providing potential contributors with valuable information. They provide detailed

information about who needs money and when and how much. By providing such information, solicitations reduce the costs of participation and may increase the probability of assent.

While charities identify their potential donor pool, there is also an element of randomization because from this pool, they solicit a subset of individuals randomly. For instance, in Jonker, Paap and Franses (2000), 5,300 individuals are selected randomly out of 800,000 individuals from a large charitable organization's database in the Netherlands. Similarly, in field-work conducted by Huck and Rasul (2011) direct mail solicitations for a charitable cause were sent to individuals. The individuals were randomly selected from a database of persons who had purchased at least one ticket to attend either the opera or ballet in the 12 months prior to the mail-out. In our context, randomization is more likely to be the issue, since we consider solicitation in the previous two weeks.

Charities might target happy and healthy individuals for volunteer work since such activities often require physical and psychological effort. However, for the solicitation of donations, there is no explicit targeting based on happiness or health. While income is targeted, evidence suggests that money does not necessarily lead to greater happiness after a certain threshold is reached (Kahneman & Deaton, 2010). Therefore, we do not expect psychological better to be selected by income targeted. Even so, in the regression model, I do control for indicators of income.

For the identification of the effects of donating through IV approach, exclusion restrictions means that the instruments for donation should have no direct effect on subjective health. The nature of some of the instruments raises concerns about the validity of the assumption. For instance, one concern is that, a person has to be healthy enough to go out for coming across to a street collection. To address this concern, I control for experience of physical diseases. Moreover, the GINP survey is conducted through internet. Therefore, the respondents of this survey should be able to use computer and internet. If their physical or mental health would be so bad, they would not be able to answer the questions at the first place.

The exclusion restrictions cannot be tested formally. But if solicitation is randomly conducted, then it should be unrelated to the respondent's personal characteristics. Often randomization is not possible. Nevertheless, we can consider whether observables, $X$are balanced across differences in Z. If they are, this could justify our use of IVs. Accordingly, we check whether the observable covariates

are similar among solicited and unsolicited. There could still be concern that unobserved characteristics are related to Z. Therefore, one needs to make a decision whether the set of instruments are successful in leading to treatment that is "as good as random" or not.

Moreover, solicitation may not affect everybody in the same way. For instance, generous people who have more inclination to be pro-social might donate more than less generous people in response to a solicitation. That is, there could be heterogeneous effects. However, in model (2) we propose a linear model where the effect of solicitation on pro-social behavior is assumed to be constant. Even if the treatment effects are heterogeneous, Imbens and Angrist (1994) show that under the "monotonicity" assumption, IV estimates remain valid. The monotonicity assumption requires that the solicitation either has no effect on pro-social behavior or it influences pro-social behavior in the same direction whenever it has any impact. For instance, if people get disturbed by solicitation and give less than her intended amount, monotonicity assumption is violated.

Under these assumptions, IV results yield the local average treatment effect (LATE) for donors who donated because they are asked to do so, but would not donate if they were not asked.

4.4 Data

We employ two datasets for the empirical analysis. The first is the Longitudinal Internet Studies for the Social Sciences (LISS) panel of Centerdata; the second is the Giving in the Netherlands Panel (GNIP).

We use five waves of individual-level data from the LISS collected annually between 2007 and 2012. The LISS contains 5000 households, comprising 8000 individuals. It is based on a true probability sample of households drawn from the population register by Statistics Netherlands. As its name suggests, it is internet based panel. However, households that cannnot otherwise participate are provided with a computer and Internet connection. LISS panel members get an incentive payment of € 7.5 per half-hour of interview time.

The respondent is asked whether s/he has performed voluntary work or donated money for many different types of organizations, including a sports/outdoor activities club, a cultural association or hobby club, an organization for humanitarian aid, an organization for environmental protection, a

peace organization, animal rights organization etc. These questions are used to construct volunteer and donated dummy variables. The question for measuring volunteering time is the following: "Considered all together, how much time do you spend on voluntary work per week, on average? (Including hours that you possibly spend on informal care)". For measuring the amount of donation, the question asked in 2009 and 2010 is the following: "For donations & gifts (for family, friends, charity etc), please indicate how many Euros you spend on this personally per month, on average (considering last 12 months)". Although questions from which donated and volunteer dummy variables are derived are administered in all 5 waves, amount of donation is only available for two waves. Therefore, the number of observations for donation amount is lower.

In addition to questions on volunteering and donations, LISS respondents are asked their overall happiness. As our dependent variable, we use the respondents' happiness, derived from the following question: "On the whole, how happy would you say you are?" with 11 possible categories ranging from 0 "totally unhappy" to 10 "totally happy". The LISS panel has many variables that can be controlled for. Personal status related covariates include self-reported health status, having a long-standing disease, days hospitalized, as well as a set of socio-economic status related variables such as gender, age, urbanity of residence, education level, labor market status, monthly net income, number of household members, number of children living at home, marital status, whether the respondent is a member of social organization, and whether the respondent has participated an event in a social organization[39].

From the full sample of (N=29,264[40]), we use a subset of 22,560[41] person years, although the number of observations varies depending on the availability of data for the control variable. We drop observations from the sample if the volunteer or donation status information is not available (1,753 cases). Self reported health is not available in 3,043 cases; long-standing illness is not available in 15 cases; days hospitalized is not available in 71 cases; marital status are not available in 554 cases; place

[39] The LISS panel contains many variables on personality which may be important for donating behavior. We did not use these personality variables in our analysis because longitudinal nature of the data allows for controlling time-invariant personality factors.
[40] n= 9887, T=5 and the panel is unbalanced.
[41] n= 7564, T=5 and the panel is unbalanced.

of residence not available in 22 cases, income is missing or the respondent does not want to reply in 1,213 cases; education is missing in 33 cases.

We also use four waves of individual-level data out of five waves from the Giving in the Netherlands Panel collected biannually since 2004 because the 2002 wave does not have subjective health measure. This dataset has been used as the basis for macro-economic estimates of giving in the Netherlands. The GINP is a random sample of individuals from pool of 40,000 households available to NIPO (Netherlands Institute of Public Opinion). The pool is representative of the Dutch population with respect to gender, age, level of education, home ownership, household size and region. Completing a survey is rewarded with a number of token points depending on the length of the survey completed. At the end of the survey, the points can be exchanged for a voucher or alternatively, for a donation to a charitable cause.

In May-June 2002, a representative sample of 1,707 individuals completed the GINPS. The baseline sample was followed up in May of 2004, 2006, 2008 and 2010. There is both entry into and exit from the panel, leading to unbalanced data with an increasing number of individual interviews over time. Unfortunately, a measure of subjective wellbeing is available only for the 2006 wave. Since a subjective health measure is available in four waves from 2004 to 2010, we also used subjective health as a dependent variable. The question for measuring subjective health is "What do you think about your health in general?" with five possible categories ranging from 1, "Bad," to 5, "Excellent".

Respondents are asked to report whether they donated and the amount donated as a result of 13 different types of fundraising efforts in the previous two weeks. The total amount of the donation is calculated based on answers given to these questions. The amount donated is not asked in the 2004 wave, only whether the respondent has donated or not is asked in 2004 wave, therefore the number of observations for the amount donated is lower.

Questions on volunteering are asked in all waves. Respondents are asked whether they have performed voluntary work in various fields of services such as sports, health, social care provision, education, etc. This 0/1 type of question is used to construct volunteer dummy variables. The

respondents are also asked the following question: "How many hours a month did you normally spend to volunteering (reference last year)?"

The GINP has fewer covariates. The control variables for explaining life satisfaction and self-reported health are gender, age, community size, province lived, education level, labor market status, number of household members, number of children in the household, marital status, and income from employment.

Of the full sample of (N=6,421)[42], the analysis is based on 4,847[43] observations. The observations for which donation status (437 cases) and volunteering status (482 cases) information are not available are dropped. Observations are dropped if education (7 cases), household size (1 case) number of children (1 case) province (2 cases) community size (2 case) are not available. We also dropped observations for which income is missing or where the respondent prefers not to say (642 cases).

The analysis from the LISS uses 22,560 observations from 7,564 individuals. Descriptive statistics appear in Appendix Table A-1. Overall, 39 % of the respondents say that they donated to a social organization. The average amount of donation for respondents who donated is € 45.71. Happiness scores are skewed towards the top of the possible answer distribution, with over 90% of the sample scoring more than 5 on the happiness scale.

The GINP analysis uses 4,847 observations from 2,740 individuals. Descriptive statistics appear in Appendix Table A-2. Overall, 44% of respondents say they donated in the last two weeks; the average donation is € 5.37 and the average donation among donors is €12.14. Subjective health scores are skewed towards the top of the possible answer distribution with over 80% responding good, very good or excellent.

The LISS panel is a true probability sample of the households in the Netherlands whereas GINP is an opt-in web-panel. Although the pool where GINP respondents were drawn is representative of Dutch population with respect to gender, age, level of education, home ownership, household size and region, there might still be self-selection. Given that there is growing evidence that opt-in web-based

[42] n=3,367; T=4, and the panel is unbalanced.
[43] n=2,740; T=4, and the panel is unbalanced.

research is not as accurate as research using probability sampling (see especially Yeager et al., 2011), we should avoid making claims about effect sizes and generalizability of the finding in a population from the GINP analysis. Yet, this does not mean that nonprobability samples have no value. We can still document that our variables of interest relate which is sufficient for suggesting the existence of warm glow.

Annual donation from the GINP could be estimated as around € 120 (= €5.37*24). When we compare the donation amounts in the LISS panel to the GINP, there is a large discrepancy. Since the LISS panel asks the respondents for an annual estimation, there might be a large recall bias in the estimate. However, the GINP asks respondents how much they have donated in the previous two weeks. This could potentially reduce recall bias considerably. Another check is on the timing of the fieldwork. The LISS panel collects pro-social behavior indicators in February whereas the GINP conducts the fieldwork in May. According to Centraal Bureau Fondsenwerving's national schedule on door to door fundraising, neither February nor May is particularly crowded with door to door fundraising. (CBF, 2013) On the other hand, it is also possible that socially engaged people are more likely to participate in an opt-in panel, so GINP may overestimate the amount donated.

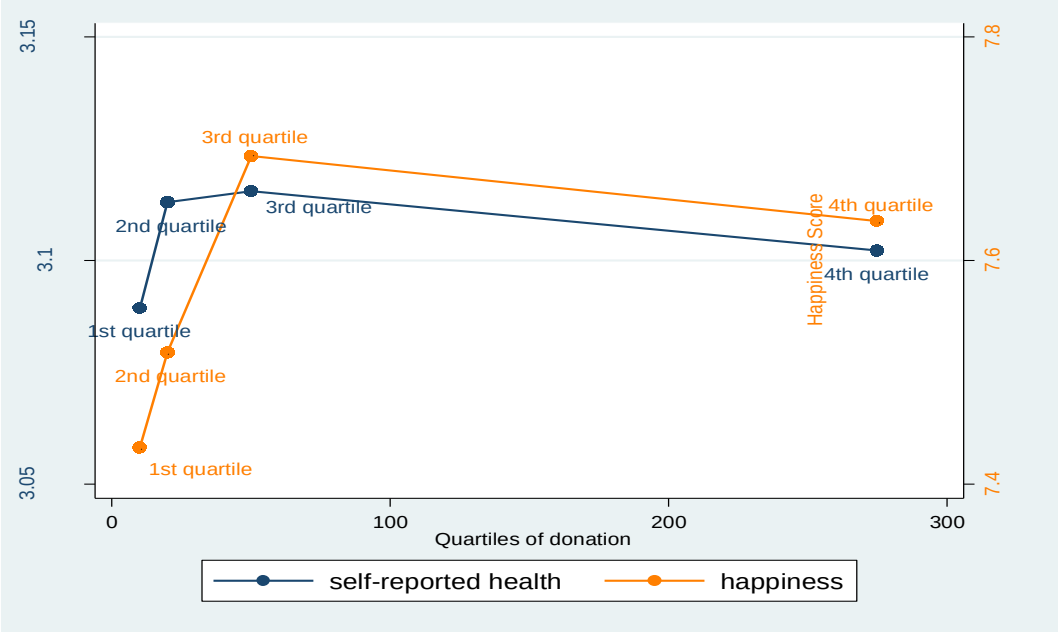
Another disadvantage of the LISS panel is that pro-social behavior and happiness are not measured simultaneously; by contrast, the GINP measures life satisfaction, self-reported health and pro-social behavior simultaneously. An advantage of both surveys is that because they are internet based, respondents are less likely to try to present a desirable social image. However, since the GINP asks about making a "donation in the last two weeks", this recency could prompt respondents to reproduce a desirable social self-image. Also, fixed effects models deal with respondents' tendency to overstate their donating and volunteering.

4.5 Proxies for Subjective Wellbeing

We use happiness, life satisfaction and a subjective health measure as proxies for subjective wellbeing. There is no controversy in the literature for using happiness and life satisfaction as indicators of subjective wellbeing. But, although subjective health contains psychological health and and answers to that question might reflect the mood of the respondent, it is a crude measure with a

strong possibility of error, partly because a subjective health measure also contains objective health information. Therefore, we want to determine whether subjective health is a reasonable measure subjective wellbeing. A first argument in defense of self-reported health measures comes from evidence that self-reported health has 0.29 correlation with happiness score in the LISS panel and 0.27 correlation with life-satisfaction score in the GINP.

Figure 4-1: Self-reported Health and Happiness by Quartiles of the Amount Donated



Source: LISS Panel, 2009-2010

Figure 4-1 shows happiness and subjective health for each quartile of donation amount using LISS data. As the figure indicates, happiness and subjective health follow similar patterns for donors. Based on our findings, we conclude that subjective health is a reasonable measure of subjective wellbeing.

4.6 Empirical Results

In the regression models, the happiness score is treated as continuous. All the models report the contribution of income to subjective wellbeing measures so that the reader can put the size of the effect into the context by comparing the coefficient of income and indicators of donating. For studying the impact of donating, regression models with different specifications were used. In the first column, amount donated is included in the model whereas the second column includes dummy variables for donating to capture potential non-linear relationships between happiness and pro-social

behavior. To check different types of non-linear relationships, in the third column, amount donated and dummy variable for donated is included whereas in the fourth column, amount donated and its squared is included. All the following regression results follow the same structure.

4.6.1 LISS Panel

Table 4-1 shows the random effects regression estimates of the effects of pro-social behavior on happiness. In the first column, the amount donated is not statistically significantly related to happiness; in the second column, donated dummy variable is positively and significantly associated with happiness. These results suggest some sort of non-linear relationship between pro-social behavior and happiness. In the third column, amount donated and dummy variable for donating are not statistically significant at 10% level. The fourth column suggests a concave relationship between amount donated and happiness.

One concern with random-effects results is the presence of unobserved differences in individual innate characteristics. Personality traits such as extroversion could be important for happiness, and certain characteristics may simultaneously promote happiness and pro-social behavior. Given that personality might jointly influence happiness and pro-social behavior, random effects results may be unreliable since these innate characteristics are not adequately controlled for in these regressions.

Table 4-1: Random Effects Regressions Output for Happiness

	(1)	(2)	(3)	(4)
Income	0.00279**	0.00086	0.00242**	0.00247**
(in € 1000)	(2.50)	(0.73)	(2.22)	(2.29)
Amount donated	0.23848		0.22595	0.85361***
(in € 1000)	(1.27)		(1.21)	(2.78)
Amount donated squared				-0.89378**
(in € 1000)				(-2.40)
Donated dummy		0.05779***	0.04365*	
		(3.97)	(1.76)	
Baseline Controls	+	+	+	+
N of Observations	7,697	22,560	7,697	7,697
Number of id	4,822	7,564	4,822	4,822
Joint F test p value	0.20	0.00	0.11	0.02
Hausman Test p value	0.00	0.00	0.00	0.00

Std. errors are robust, clustered at individual level, z-statistics in ( ) *** p<0.01, ** p<0.05, * p<0.10

The longitudinal structure of the LISS panel allows controlling for unobserved individual differences with a fixed effects model. However, the considerable number of respondents already

donated (39% donated money). For fixed effects model to be identified, we need transitions[44] in donating status over time. In the LISS panel, we see that 61.8% of those who donated in one period also donated in the next period, while 79% of those who did not donate in one period did not donate in the next. We also see considerable variation in donating status: 21% of those who did not donate in one period changed their behavior in the next; conversely, 38% of those who donated in the one period did not donate in the next.

The results of fixed effects models are provided in the Table 4-2. The amount donated is not statistically significant in the first column whereas, in the second column, dummy variable for donating is statistically significant. In the third and fourth column, we do not detect any statistically significant relationship between indicators of pro-social behavior and happiness.

The fixed effects results are interesting; the effect of engaging in donation remains after controlling for extensive set of control variables which possibly constitute major shocks to happiness such as changes in marital status, income, employment etc. Comparing the coefficient of income and donated dummy from column 2 in Table 4-2 suggests that engaging in donating brings much larger happiness than increases in income.

Table 4-2: Fixed Effects Regressions Output for Happiness

	(1)	(2)	(3)	(4)
Income	-0.00202	0.00073	-0.00212	-0.00210
(in € 1000)	(-1.03)	(0.37)	(-1.03)	(-1.03)
Amount donated	-0.16473		-0.15266	0.08736
(in € 1000)	(-0.91)		(-0.84)	(0.25)
Amount donated squared				-0.33775
(in € 1000)				(-0.93)
Donated dummy		0.04501***	0.02467	
		(2.77)	(0.73)	
Baseline Controls	+	+	+	+
N of Observations	7,697	22,560	7,697	7,697
Number of id	4,822	7,564	4,822	4,822
Joint F test p value	0.36	0.01	0.48	0.29

Std. errors are robust, clustered at individual level, z-statistics in ( ) *** p<0.01, ** p<0.05, * p<0.10

To decide whether random effect results are justifiable, we conduct the Hausman test. Since all p values reported in the last row of Table 4-1 are 0.00, fixed effects results are preferred. From the

[44] We assume these transitions are exogenous. Yet, even if it is driven by income shocks, it does not pose a problem for our estimation methodology since income is included in the set of control variables.

LISS data, we do not find the amount donated to have any effect on happiness, but engaging in pro-donation is associated with higher happiness scores (in column 2 of Table 4-2).

4.6.2 Giving in the Netherlands Panel

All the regression results presented in this section follow the same structure as described in the previous part. Table 4-3 provides random effects estimates of the effect of pro-social behavior on self-reported health. Here, the self-reported health score is again treated as continuous. Neither the amount donated nor dummy variable for donating is statistically significantly related to subjective health in the random effects specification.

Table 4-3: Random Effects Regressions Output for Self-Reported Health

	(1)	(2)	(3)	(4)
Income	0.05251***	0.04698**	0.05249***	0.05258***
(in € 1000)	(2.75)	(2.46)	(2.75)	(2.76)
Amount donated	0.08481		0.07305	-0.06756
(in € 1000)	(0.40)		(0.34)	(-0.10)
Amount donated squared				0.14313
(in € 1000)				(0.32)
Donated dummy		0.00360	0.00503	
		(0.16)	(0.22)	
Baseline Controls	+	+	+	+
Observations	4,631	4,847	4,631	4,631
Number of id	2,655	2,740	2,655	2,655
Joint F test p value	0.69	0.87	0.91	0.02[45]
Hausman Test p value	0.00	0.00	0.00	0.00

Std. errors are robust, clustered at individual level, z-statistics in ( ) *** p<0.01, ** p<0.05, *p<0.10

Table 4-4: Fixed Effects Regressions Output for Self-Reported Health

	(1)	(2)	(3)	(4)
Income	0.03400	0.03759	0.03391	0.03420
(in € 1000)	(1.03)	(1.14)	(1.03)	(1.03)
Amount donated	0.25044		0.27576	-0.38948
(in € 1000)	(1.12)		(1.24)	(-0.47)
Amount donated squared				0.55449
(in € 1000)				(1.01)
Donated dummy		-0.00916	-0.01018	
		(-0.34)	(-0.36)	
Baseline Controls	+	+	+	+
Observations	4,631	4,847	4,631	4,631
Number of id	2,655	2,740	2,655	2,655
Joint F test p value	0.26	0.74	0.45	0.00[46]

Std. errors are robust, clustered at individual level, z-statistics in ( ) *** p<0.01, ** p<0.05, * p<0.10

[45] Without clustering Joint F test p value is 0.93
[46] Without clustering Joint F test p value is 0.53.

Table 4-4 provides fixed effects estimates of the effect of pro-social behavior on self-reported health. Similar to random effects results, we do not detect any statistically significant relationship between indicators of pro-social behavior and self-reported health in any of the four models. To decide whether random effect results or fixed effects results are justifiable, we conduct a Hausman test. Since all p values reported in the last row of Table 4-3 are 0.00, the fixed effects results are more reliable.

Table 4-5 reports OLS estimates of the effect of pro-social behavior on life satisfaction using the 2006 wave of the GINP. Unlike the LISS panel results, in the last two columns where life satisfaction is the dependent variable, indicators for donation are not statistically significant in the first two columns. The difference might stem from the fact that the LISS looks at annual donations whereas the GINP looks at donating in the previous two weeks. Like the LISS random effects results, the fourth column suggests a concave relationship between amount donated and life satisfaction.

Table 4-5: OLS Regressions Output for Life Satisfaction (2006 only)

	(1)	(2)	(3)	(4)
Income	0.12760***	0.13301***	0.13291***	0.13091***
(in € 1000)	(2.60)	(2.74)	(2.73)	(2.68)
Amount donated	0.67809		0.35825	8.14134*
(in € 1000)	(0.31)		(0.15)	(1.78)
Amount donated squared				-59.21891**
(in € 1000)				(-2.39)
Donated dummy		0.00536	0.00185	
		(0.08)	(0.02)	
Baseline Controls	+	+	+	+
Observations	1,305	1,305	1,305	1,305
Joint F test p value	0.76	0.94	0.98	0.02

Std. errors are robust, z-statistics in ( ) *** p<0.01, ** p<0.05, * p<0.10

From the GINP, overall, we do not observe any statistically significant relationship between pro-social behavior and self-reported health whereas we do see a concave relationship between amount donated and life-satisfaction.

4.6.3 IV Estimates

Potentially, there could be endogeneity issues for donating, even for the fixed effects results. Shocks to happiness might be driving the results that we observe rather than shocks to donation increasing happiness. For instance, if a person gets happier, she might want to share it with others in the form of higher pro-social behavior. Given this reverse-causality is a concern for our fixed effects

results, therefore; in this section, we use instruments for donating. As mentioned in Section 3, we first discuss the validity of our instruments.

Table 4-6: First stage regressions of Donation and Donated Dummy

	Donation	Donation	Donated	Donated
	(OLS)	(OLS)	(Probit)	(Probit)
Door to door collection	0.074	2.727**	0.746***	1.460***
	(0.06)	(2.46)	(13.30)	(11.63)
Street Collection	-0.447	-1.078	-0.069	-0.168
	(-0.28)	(-0.59)	(-0.95)	(-0.77)
Sponsor Campaign	4.631***	1.905	0.137*	0.301*
	(2.75)	(1.28)	(1.74)	(1.77)
Collection in the Church	14.173***	9.696***	0.828***	1.678***
	(9.98)	(8.04)	(11.03)	(8.13)
Collection at work	2.635	0.434	0.337	0.893*
	(0.60)	(0.10)	(1.56)	(1.69)
Television campaign	-1.610	2.261*	-0.431***	-0.881***
	(-0.94)	(1.85)	(-5.49)	(-6.20)
Direct mail letter	3.598***	5.809***	-0.091	0.071
	(2.67)	(5.06)	(-1.44)	(0.54)
Via internet/e-mail	-2.390	-1.767	-0.131	-0.544**
	(-0.82)	(-0.63)	(-1.04)	(-2.12)
Collection during an event	-2.427	4.116	-0.044	0.804
	(-0.68)	(0.89)	(-0.27)	(1.52)
Collection via membership organization	1.396	6.589	0.229	0.333
	(0.34)	(1.36)	(1.11)	(0.65)
Appeal via advertisement	3.938	-5.654**	-0.027	-0.254
	(1.33)	(-2.21)	(-0.20)	(-1.05)
Via buying something	3.975**	0.764	0.004	0.127
	(2.00)	(0.44)	(0.05)	(0.55)
Via lottery tickets	1.849	8.598***	0.001	0.470**
	(1.19)	(5.37)	(0.01)	(2.31)
Was not asked to donate	-2.867**	-0.744	-1.769***	-47
	(-2.04)	(-0.60)	(-21.77)	
Baseline controls	+	+	+	+
Dummies for Diseases	-	+	-	+
Test of Ho: the instruments are jointly 0				
F statistic (Chi2)	15.59	17.23	1481.31	215.51
P value	0.0000	0.0000	0.0000	0.0000
N	4,631	1,305	4,847	994

[47] No solicitation predicts failure perfectly, therefore 527 observations are dropped.

Table 4-6 shows the estimates of the first stage of regression. Column (1) corresponds to a specification that controls for all control variables, X. The results indicate that solicitation is an important determinant of donation. As discussed earlier, F statistic above roughly 10 makes IV inference sufficiently strong. The F statistics presented in the last row of each column clearly exceed 10 and p values are 0.00; we therefore conclude that we do not have a weak instrument problem. When dummy variables for various diseases are added to control for physical health, the F statistic remains almost the same (shown in column 2). This is reassuring for the use of our instruments.

Exclusion restrictions cannot be tested formally. To check the plausibility of our instruments, we firstly check whether the observable characteristics of the individuals are balanced over our instrument set. That is, we ask to ourselves, in terms of observable characteristics, are individuals who are more exposed to solicitation similar to individuals who are less exposed to solicitation. In the IV analysis, we use 13 different types of solicitation. For ease of exposition, we document whether people exposed to above median number of solicitations are different from people exposed to below median solicitations. Table A-3 provides sample means of all controls used in the regression analysis by median solicitation status. There is some difference between those who are more solicited and those who are less solicited in terms of age, gender, marital status, province of residence, and community size. As expected, those who are more solicited have higher income from employment, yet the difference between the two columns is not statistically significant. The economics literature notes a health gradient of income, and this might be a problem for interpreting self-reported health. But income does not tell the whole story. Predicted health scores for people exposed to above median solicitations and below median solicitations can be compared to get an overall sense of the direction of selection. We have regressed all control variables on self-reported health and obtain predicted values using FE, RE and OLS (see bottom lines of Table A-3). The average predicted health scores for above median people are lower than those for below median people no matter which specification is used. So charities' targeting based on income does not appear to select healthier people.

Given the proliferation of fundraising activities by charities, unsolicited people are not likely to be a large group. Nonetheless, Table A-4 indicates, 39% of the respondents were not solicited in the last two weeks. We also check whether there is any transition for solicitation status over time. That is,

whether the same people are or are not solicited in each wave. We find that 68% of those who are solicited in one period are solicited in the next; 32% (521 cases) of those solicited in one period are not solicited in the next. Similarly, 52% of those who are not solicited in one period are not solicited in the next period; 48 % (365 cases) who are not solicited in one period are solicited in the next. Therefore, there is considerable change in respondents' experiences of solicitation.

Since we have 13 instruments, we can check over-identifying restrictions, which gives some idea on the validity of our instruments. One note of caution here is that the test assumes that at least one instrument is valid. As the last row of first two columns in Table 4-7 show, our instruments jointly pass the over-identification test at 5% significance level which provides suggestive evidence for the exogeneity of our instruments.

The monotonicity or "no defiers" assumption is another concern. For the purposes of our study, monotonicity implies that with one more solicitation, the respondent might be induced to donate, but should not give less than her intended amount. In fact, some people might feel discouraged when they are solicited. Diamond and Noble (2001) find that in response to frequent solicitations, donors develop defense mechanisms, for example, simply throwing away mail requests. Using data from some large Dutch charities' database, Van Diepen, Donkers and Franses (2006) find that additional appeals initially generate more donations, but at a certain point, donors become irritated. Like exclusion restrictions, the monotonicity assumption cannot be tested formally. However, we can convert different types of solicitation into numbers of solicitations[48] to study potential discouragement at higher levels of solicitation. Table A-5 reports regression results for donation by different numbers of solicitations compared to a reference category of no solicitation. In column 1, all coefficients are positive. In column 2, except for nine solicitations, again all coefficients are positive. In columns 3 and 4, all coefficients are positive and increasing as well. The results suggest that respondents are not less likely to give as a result of increased solicitation. In sum, we find no evidence that the monotonicity assumption is implausible.

Table 4-7 presents fixed effects IV regression results for self-reported health from the GINP. All results include full control variables. We also report the overidentifying restrictions test's (Hansen J

[48] Table A-4 in Appendix shows distribution of # of solicitations.

statistic) p value and a test of exogeneity (Davidson-MacKinnon test) of pro-social behavior. The exogeneity test p value in the first column suggests that the amount donated is endogenous for explaining self-reported health. Overidentfying restrictions test p values suggests that our instrument set is reasonably valid. Contrary to intuition, in columns 1 and 4, we find a negative impact of the amount donated on subjective health, but the coefficient is significant at 10% level.

Table 4-7: Fixed Effects IV Regression Output for Self-Reported Health, GINP

	(1)	(2)	(3)	(4)
Income	0.04159	0.03707	0.04160	0.04180
(in € 1000)	(1.23)	(1.13)	(1.22)	(1.24)
Amount donated	-6.32224*		-6.24935	-6.33451*
(in € 1000)	(-1.79)		(-1.47)	(-1.69)
Amount donated squared				-0.07129
(in € 1000)				(-0.01)
Donated dummy		-0.05125	-0.00243	
		(-1.10)	(-0.04)	
Baseline Controls	+	+	+	+
N of Observations	4,631	4,847	4,631	4,631
Number of id	2,655	2,740	2,655	2,655
Overid. Restr. P value	0.79	0.41	0.72	0.72
Test of exogeneity (p value)	0.04	0.27	0.11	0.15
Joint F test p value	0.07	0.27	0.19	0.19

Std. errors are robust, clustered at individual level, z-statistics in ( ) *** p<0.01, ** p<0.05, * p<0.10

Table 4-8 shows IV regression results for life satisfaction. All results include full control variables. We also report the overidentifying restrictions test's (Sargan's statistic) p value and a test of exogeneity (Wu-Hausman statistic) for pro-social behavior. In the first, third and fourth models, exogeneity test being rejected suggest that pro-social behavior is endogenous for explaining life satisfaction. Given that overidentifying restrictions test's p value is lower than 0.05 in the first, third and the fourth column, we conclude that our instruments are reasonably valid. The IV estimates from model (1) and model (3) suggest the amount donated influences life satisfaction. However, we find no effect of engaging in pro-social behavior on life satisfaction.

The IV results from model (1) suggest that increasing one's donations by €1 increases life satisfaction by 0.013 points on the 0-7 life satisfaction scale. To make sense of the results, we compare the coefficient of income and donation. Column 1 suggests that for equivalent life

satisfaction, increasing one's donations by €1 corresponds to a €104[49] increase in one's employment income. These estimates show the effect of donating on life satisfaction for "compliers", i.e., the effect of donating on life satisfaction for people who donate € 1 more because they were asked to do so, but would not donate if they were not asked. Whether we can generalize the effects of donating on life satisfaction to a whole population depends on the characteristics of "compliers". This group of individuals is not likely to be representative of the Dutch population. Thus, IV estimates might not reflect the average treatment effect. Therefore, extrapolation is not meaningful. This €104 is put to give a feeling about the effect size.

Table 4-8: IV Regression Output for Life Satisfaction, GINP

	(1)	(2)	(3)	(4)
Income	0.12226**	0.13301***	0.11955**	0.12345**
(in € 1000)	(2.47)	(2.77)	(2.39)	(2.49)
Amount donated	12.65022**		19.31787**	5.02620
(in € 1000)	(2.03)		(2.03)	(0.40)
Amount donated squared				95.98539
(in € 1000)				(0.56)
Donated dummy		0.00577	-0.14116	
		(0.07)	(-1.20)	
Baseline Controls	+	+	+	+
N of Observations	1,305	1,305	1,305	1,305
Overid. Restr. P value	0.27	0.09	0.35	0.22
Test of exogeneity (p value)	0.1	0.99	0.02	0.09
Joint F test p value	0.04	0.94	0.11	0.23

Std. errors are robust, z-statistics in ( ) *** p<0.01, ** p<0.05, * p<0.10

4.7 Robustness Checks

This section examines the robustness of our baseline IV results to different specifications, selection of analysis sample, and selection of instruments. The results appear in Table 4-9. We first control for whether the respondent or anyone close to her experienced various diseases in last 12 months in our regression model. Eleven dummy variables are generated for 11 diseases.[50] Since subjective health is a crude measure of wellbeing, including various diseases better accounts for physical health. One limitation with this approach is that the question about medical problems in last 12 months was administered only in 2004 and 2006. Therefore, the sample size is smaller and

[49] 95% confidence interval of the estimate is (-34.19, 241.13) and 90% confidence interval of the estimate is (-12.06, 218.99)

[50] The diseases are bronchial, cardiovascular, stomach, liver, intestine, kidney, joints, diabetes, nervous system, skin, cancer, brain and other diseases.

standard errors are higher than in the baseline IV estimates. Including disease dummy variables renders the effect of amount donated on subjective health insignificant. For column 2, including disease dummy variables does not change the magnitude of the effect of amount donated on life satisfaction.

Table 4-9: Robustness Checks, GINP

	Self-Reported Health	Life Satisfaction
A. Adding Disease controls		
Amount Donated (in € 1000)	-3.54097 (-0.90)	12.65195** (2.13)
Observations	1,305	1,305
B. Using 2006 wave only		
Amount Donated (in € 1000)	-2.54495 (-0.59)	12.65022** (2.03)
Observations	1,305	1,305
C. Using household income		
Amount Donated (in € 1000)	-8.44087* (-1.91)	13.95156** (2.01)
Observations	3,780	1,104
D. Using a subset of instruments		
Amount Donated (in € 1000)	-7.61797* (-1.83)	12.53001* (1.88)
Observations	4,631	1,305

Std. errors are robust, self-reported health results are clustered at individual level, life satisfaction results are not clustered, z-statistics in ( ) *** p<0.01, ** p<0.05, * p<0.10

In panel B, to make the results in Table 4-7 and 4-8 comparable, we run model 1 of Table 4-7 using 2006 wave data only. The relationship between amount donated and self-reported health becomes insignificant when 2006 wave data is used. Based on checks from Panel A and B, the results displayed in Table 4-7 are not robust to inclusion of physical diseases and using 2006 wave data. Therefore, we conclude that subjective health is not a good measure of subjective wellbeing.

In the baseline model, we use income from employment since it has no categories and it allows for marginal effect comparisons. However, it does not reflect household income fully. We have household income variable but it is bracketed by €500 increments. We therefore check whether using household income rather than income from employment makes any difference. Panel C reports very similar estimates to our baseline results.

Panel D checks the sensitivity of our baseline results to the different subset of instruments. Our baseline results reports IV estimates using all of the variables in the instrument set. However, charities

might easily target through personal soliciting means such as direct post-mail, e-mail, door to door collection. Moreover, being solicited through a collection via membership organization might capture sociability differences between people since sociable people are more likely to be members of such organizations. We discard direct personal solicitations and collection via membership organizations and restrict our instrument set to the following; appeal via street collection, sponsor campaign, church collection, collection at work, television campaigns, an event, advertisement, buying something, lottery tickets, or not solicited. The results are close to our baseline results but slightly less precise.

4.8 Heterogeneous effects

OLS yields the average effect of donating on subjective wellbeing whereas IV estimates yield the effect for "compliers". In our context, "compliers" are those whose donation status is affected by solicitation. IV estimates in the baseline model are the weighted average of effects in each subgroup. That is, if the solicitations are very strong predictors of donating within a specific sub-group, that subgroup weight is higher in LATE calculation; if solicitations are not related to donations for another subgroup, that subgroup is not reflected in IV estimates. In this section, we check the effects for different groups to determine if there are qualitative differences between subgroups. The results are reported in Table 4-10.

Panel A shows the difference in the effects for males and females. Column 1 reports the IV estimates of the relationship between the amount donated and self-reported health and column 2 reports IV estimates of the relationship between the amount donated and life-satisfaction. The IV estimate for females is much higher than for males (6.86 versus -9.96), but the coefficients for both are statistically insignificant. Similarly, column 2 suggests that males are less likely to have increased life satisfaction as a result of donating extra euro (.006) than females (.011). However, the coefficient of amount donated variable is significant in none of the models.

Panel B present IV estimates for the sample disaggregated by median age: younger than 45, older than 45. The effect of donating an extra euro on self-reported health is negative but insignificant both for people older than 45 and younger than 45. Column 2 shows the IV estimates for life satisfaction, indicating that donating an extra euro impacts life satisfaction negatively for those below the median

age and positively for those above the median age. The effect for those above the median age is larger, and more precisely estimated.

Panel C show the effects of donating an extra euro decomposed by median income from employment. The effect of donating an extra euro on subjective health is much lower in magnitude for people above median income compared to people below median income. Yet, none of them are statistically significant. The effect of donating an extra euro on life satisfaction is much higher in the below median income group than the above median income group (.045 vs. -.017). However, there is no statistically discernible difference between below median income and higher median income in terms of the effect of donating an extra euro. (see column 2).

Table 4-10: Heterogeneous Effects, GINP

	Self-Reported Health	Life Satisfaction
A. Gender		
Amount Donated	-9.95938	6.18902
(in € 1000)	(-1.17)	(0.30)
Amount Donated*Female	6.85975	11.36736
	(0.50)	(0.37)
N of Observations	4,631	1,305
B. Age<= median(45)		
Amount Donated	-6.63315	24.68398*
(in € 1000)	(-1.52)	(1.88)
Amount Donated*	-4.02980	-37.66508
(Age <=Median (45))	(-0.36)	(-1.06)
N of Observations	4,631	1,305
C. Employment Income <= median(450 Euro)		
Amount Donated	-8.11978	-16.53513
(in € 1000)	(-1.09)	(-0.58)
Amount Donated*	2.78928	45.19118
(Emp. Income<=Median (450 Euro))	(0.27)	(1.01)
N of Observations	4,631	1,305
D. Volunteered		
Amount Donated	-8.55449	10.18665
(in € 1000)	(-1.13)	(0.50)
Amount Donated* Volunteer	4.20032	-1.06298
	(0.35)	(-0.03)
N of Observations	4,631	1,305

Std. errors are robust, self-reported health results are clustered at individual level, life satisfaction results are not clustered, z-statistics in ( ) *** p<0.01, ** p<0.05, * p<0.10

Sociability might be one dimension where heterogeneous effects could be observed. In our context, our best proxy for sociability is volunteering since volunteers are generally the ones who are members of some social organization. Panel D presents IV estimates of the effects of donating an extra euro disaggregated by volunteering status. In column 1, we look at subjective health. Donating

has a negative effect on subjective health for non-volunteers, but a positive effect on volunteers. Yet, both are imprecisely estimated. Column 2 suggests that donating increases life-satisfaction for volunteers, but it decreases life-satisfaction of non-volunteers. For both volunteers and non-volunteers, the effects are statistically insignificant.

Results from Table 4-10 suggest that largest effects of donation on life satisfaction occur for females, older than age 45, below median income, and who have not volunteered in the previous 12 months.

4.9 Conclusion

In this paper, we explore "warm glow" motivation of pro-social behavior. Although "warm glow" was first hypothesized at the end of the 1980s in economics literature, no empirical analysis has been conducted to measure of the effect of pro-social behavior on happiness in monetary terms. There is an extensive psychology literature on the relationship between happiness and pro-social behavior, but these are mainly experimental studies conducted with small groups of students and unrealistic donation amounts.

In our study, we ask whether pro-social behavior makes people happier. To identify causality, we use fixed effects specification and instrumental variables. We use different types of personal solicitation as instruments for donating since the literature suggests that solicitation is an effective way to induce people to make charitable donations. We use two datasets from the Netherlands: the first is five waves of the LISS and the second is four waves of the GINP. Our results suggest that donating increases life satisfaction. However, given the opt-in nature of GINP, the magnitude of the effect from GINP may not be extrapolated to the population.

This result could be useful for policy making. A public policy that fosters donating might increase happiness in a society since it contributes to solving problems for needy recipients while yielding psychological benefits (i.e., happiness) to the donors. In fact, some organizations seem to be using this sort of incentive in their campaigns, as opposed to the classic campaigns involving images of needy individuals designed to elicit sympathy and guilt. For instance, American Red Cross tells prospective blood donors that "the need is constant, the gratification is instant."

Yet incentivizing behaviors might have detrimental results. If the concept of benefitting from donating is introduced, this might destroy extrinsic motivations. For instance, the mere thought of money undermines people's motivation to give (Vohs, Mead, & Goode, 2006). For studying reciprocity, randomly some potential donors received large gifts from the soliciting charity and some potential donors did not receive anything. Potential donors who received large gifts are found to donate smaller amounts; those who do not receive gifts are more likely to donate larger amounts (Falk, 2007). Therefore, further research is needed for clarifying whether advertising these psychological benefits of charitable giving undermines the happiness donors receive from giving to others, crowds out intrinsic motivations, or decreases subsequent giving.

Appendix

Table A-1. Data Descriptions, Sample Means, and Standard Deviations LISS Panel: 2007-2012[51]				
Variables	Description	Mean(Std. DevAll)	(within)	Obs (Groups)
Happy	Happiness score, coded so that 0 =totally unhappy, 10 = totally happy	7.58 (1.26)	(0.64)	N = 22560 (n = 7564)
Health	Individual assessment of health in the past year; 1 = very poor, 5 = excellent	3.13 (0.75)	(0.37)	N = 22560 (n = 7564)
Long-standing disease	Suffering from a long-standing disease, affliction, handicap, or an accident	0.3 (0.46)	(0.17)	N = 22560 (n = 7564)
Days hospitalized	Time spent in hospital or a clinic over the past 12 months (days)	0.53 (4.75)	(3.67)	N = 22560 (n = 7564)
Female	Gender of the respondent, women = 1	0.54 (0.5)	(0.01)	N = 22560 (n = 7564)
Age	Age in CBS (Statistics Netherlands) categories	4.87 (1.59)	(0.22)	N = 22560 (n = 7564)
Married	Civil status, married=1	0.6 (0.49)	(0.09)	N = 22560 (n = 7564)
Never been married	Civil status, Never been married=1	0.27 (0.44)	(0.06)	N = 22560 (n = 7564)
Not urban	Urban character of place of residence, not urban=1	0.15 (0.36)	(0.05)	N = 22560 (n = 7564)
Slightly to moderately urban	Urban character of place of residence, Slightly urban=1 | Moderately urban=1	0.45 (0.5)	(0.06)	N = 22560 (n = 7564)
Employed	Paid employment| in family business| freelancer, or self-employed	0.54 (0.5)	(0.15)	N = 22560 (n = 7564)
Income	Personal net monthly income in Euros	1595.5 (5644.26)	(3040.64)	N = 22560 (n = 7564)
Education	Level of education in CBS (Statistics Netherlands) categories	3.43 (1.51)	(0.28)	N = 22560 (n = 7564)
Household size	Number of household members	2.66 (1.31)	(0.26)	N = 22560 (n = 7564)
# of children	Number of living-at-home children in the household	0.86 (1.14)	(0.23)	N = 22560 (n = 7564)
Member	Member of a social organization	0.71 (0.45)	(0.28)	N = 22560 (n = 7564)
Participated an event	Participated an event in a social organization	0.33 (0.47)	(0.34)	N = 22560 (n = 7564)
Donation amount	Monthly donation and gifts (for family, friends, charity etc) expenditure (past 12 months)	37.4 (64.1)	(35.76)	N = 7697 (n = 4822)
Donated dummy	Dummy for donating	0.39 (0.49)	(0.31)	N = 22560 (n = 7564)
Volunteering hours	Hours spent on voluntary work per week on average over past 12 months	2.04 (6.07)	(3.94)	N = 22560 (n = 7564)
Volunteer dummy	Dummy for volunteering	0.36 (0.48)	(0.26)	N = 22560 (n = 7564)

[51] Average years of participation for the sample is 2.98

Table A-2: Data Descriptions, Sample Means, and Standard Deviations GIN Panel: 2004-2010[52]				
Variables	Description	Mean (Std. Dev. All)	(within)	Observations (Groups)
Health	Individual assessment of health in the past year; 1 = bad, 5 = excellent	3.16 (0.9)	(0.34)	N =4,847 (n=2,740)
SWB	Are you satisfied with your life? 7 scale (0=No, 1=Yes)	4.31 (1.28)		N=1,305
Age	Age in years at interview	46.55 (16.29)	(1.31)	N =4,847 (n=2,740)
Female	Gender of the respondent, women = 1	0.52 (0.5)	(0.02)	N =4,847 (n=2,740)
Married	Marital status, married=1	0.68 (0.47)	(0.19)	N =4,847 (n=2,740)
Never been married	Marital status, Never been married=1	0.29 (0.45)	(0.16)	N =4,847 (n=2,740)
Household size	Number of household members	2.61 (1.26)	(0.22)	N =4,847 (n=2,740)
# of Kids	Number of children	0.84 (1.08)	(0.22)	N =4,847 (n=2,740)
Province Lived	Province of residence			
	Groningen	0.04 (0.19)	(0.02)	N =4,847 (n=2,740)
	Friesland	0.03 (0.18)	(0.02)	N =4,847 (n=2,740)
	Drenthe	0.03 (0.18)	(0.02)	N =4,847 (n=2,740)
	Overijssel	0.06 (0.24)	(0.03)	N =4,847 (n=2,740)
	Flevoland	0.03 (0.17)	(0.02)	N =4,847 (n=2,740)
	Gelderland	0.12 (0.32)	(0.04)	N =4,847 (n=2,740)
	Utrecht	0.06 (0.25)	(0.04)	N =4,847 (n=2,740)
	Noord-Holland	0.15 (0.36)	(0.03)	N =4,847 (n=2,740)
	Zuid-Holland	0.21 (0.41)	(0.04)	N =4,847 (n=2,740)
	Zeeland	0.03 (0.18)	(0.02)	N =4,847 (n=2,740)
	Noord-Brabant	0.16 (0.36)	(0.02)	N =4,847 (n=2,740)
	Limburg	0.07 (0.25)	(0.02)	N =4,847 (n=2,740)
Community Size	Type of community			

[52] Average years of participation for the sample is 1.77.

	<10,000 inhabitants	0.02 (0.13)	(0.03)	N =4,847 (n=2,740)
	10,000 – 20,000 inhabitants	0.11 (0.31)	(0.07)	N =4,847 (n=2,740)
	20,000 – 50,000 inhabitants	0.34 (0.47)	(0.1)	N =4,847 (n=2,740)
	50,000 – 100,000 inhabitants	0.22 (0.42)	(0.07)	N =4,847 (n=2,740)
	> 100.000	0.31 (0.46)	(0.06)	N =4,847 (n=2,740)
Education	Highest completed level of education categories	4.11 (1.68)	(0.33)	N =4,847 (n=2,740)
Employed	Has paid job	0.64 (0.48)	(0.15)	N =4,847 (n=2,740)
Income	Net monthly Income from employment	772.79 (944.34)	(320.5)	N =4,847 (n=2,740)
Donated dummy	Dummy for donating or not in the last two weeks	0.44 (0.5)	(0.27)	N =4,847 (n=2,740)
Donation amount	Total amount donated in last two weeks	5.27 (29.83)	(20.98)	N =4,631 (n= 2,655)
Volunteer dummy	Dummy for volunteering or not in the last year	0.43 (0.5)	(0.21)	N =4,847 (n=2,740)
Volunteering hours	Hours usually spent on volunteering in a month (reference: last year)	5.45 (14.46)	(8.08)	N =4,847 (n=2,740)

Table A-3: Balance of observables by Median Solicitation (1 solicitation)			
	> Median	<= Median	
	Mean (SD)	Mean (SD)	Difference (Std. Error)
Age	48.94 (16.6)	45.48 (16.04)	3.459 (0.505)***
Female	0.54 (0.5)	0.5 (0.5)	0.036 (0.016)**
Married	0.72 (0.45)	0.67 (0.47)	0.056 (0.014)***
Never been married	0.25 (0.43)	0.31 (0.46)	-0.053 (0.014)***
Household size	2.63 (1.27)	2.6 (1.26)	0.032 (0.039)
# of Kids	0.83 (1.1)	0.84 (1.08)	-0.002 (0.034)
Province Lived			
Groningen	0.04 (0.19)	0.04 (0.2)	-0.006 (0.006)
Friesland	0.03 (0.18)	0.03 (0.17)	0.003 (0.005)
Drenthe	0.04 (0.2)	0.03 (0.17)	0.01 (0.006)*
Overijssel	0.08 (0.27)	0.06 (0.23)	0.023 (0.008)***
Flevoland	0.02 (0.15)	0.03 (0.17)	-0.007 (0.005)
Gelderland	0.12 (0.33)	0.12 (0.32)	0.002 (0.01)
Utrecht	0.07 (0.25)	0.06 (0.24)	0.008 (0.008)
Noord-Holland	0.13 (0.34)	0.16 (0.37)	-0.03 (0.011)***
Zuid-Holland	0.22 (0.41)	0.21 (0.41)	0.009 (0.013)
Zeeland	0.04 (0.19)	0.03 (0.18)	0.006 (0.006)
Noord-Brabant	0.15 (0.35)	0.16 (0.37)	-0.013 (0.011)
Limburg	0.06 (0.24)	0.07 (0.25)	-0.005 (0.008)
Community Size			
<10,000 inhabitants	0.02 (0.15)	0.02 (0.12)	0.007 (0.004)*
10,000 – 20,000 inhabitants	0.13 (0.34)	0.1 (0.3)	0.035 (0.01)***
20,000 – 50,000 inhabitants	0.37 (0.48)	0.32 (0.47)	0.048 (0.015)***
50,000 – 100,000 inhabitants	0.21 (0.41)	0.23 (0.42)	-0.015 (0.013)
> 100.000	0.26 (0.44)	0.34 (0.47)	-0.075 (0.014)***
Education	4.22 (1.65)	4.07 (1.7)	0.15 (0.052)***
Employed	0.62 (0.49)	0.65 (0.48)	-0.032(0.015)**
Income from employment	783.4 (1003.6)	768.08 (916.93)	15.323 (29.398)
Volunteer	0.59 (0.49)	0.36 (0.48)	0.226 (0.015)***
Predicted Subjective Health (using FE)	3.03	3.2	
Predicted Subjective Health (using RE)	3.13	3.16	
Predicted Subjective Health (using OLS)	3.14	3.15	
Predicted Life Satisfaction (using OLS)	4.39	4.3	
N	1,490	3,357	

Table A-4 : Distribution of # of Solicitations, GINP	
	Percentage
No solicitation	38.77
1 solicitation	30.49
2 solicitations	14.98
3 solicitations	8.05
4 solicitations	4.25
5 solicitations	1.96
6 solicitations	0.72
7 solicitations	0.33
8 solicitations	0.14
9 solicitations	0.23
10 solicitations	0.08
N of observations	4,847

Table A-5 : Regression Results for Donation by # of solicitation				
	Donation	Donation	Donated	Donated
1 solicitation	4.69141***	4.18274***	2.25274***	3.10345***
(ref: no solicitation)	(4.46)	(4.62)	(33.72)	(19.39)
2 solicitation	7.04867***	7.99458***	2.32396***	3.30207***
	(5.26)	(6.79)	(30.45)	(18.43)
3 solicitation	10.86927***	11.08830***	2.40552***	3.31182***
	(6.41)	(7.75)	(26.41)	(16.32)
4 solicitation	15.92029***	17.78891***	2.32329***	3.18291***
	(7.11)	(7.88)	(20.90)	(12.27)
5 solicitation	18.91780***	19.31779***	2.46004***	3.36825***
	(5.94)	(6.63)	(15.55)	(9.37)
6 solicitation	18.88510***	24.67108***	2.34976***	2.97722***
	(3.71)	(5.68)	(9.83)	(6.67)
7 solicitation	8.34907		2.40950***	
	(1.10)		(6.72)	
8 solicitation	11.00450	14.98826	2.31817***	
	(0.99)	(1.55)	(4.46)	
9 solicitation	22.88525**	-2.15704	2.88922***	
	(2.57)	(-0.16)	(5.04)	
10 solicitation	41.84475***		1.57382**	
	(2.84)		(2.46)	
Full Controls	+	+	+	+
Disease Dummies	-	+	-	+
Observations	4,631	1,305	4,847	1,517

5 Bibliography

New York Organ Donor Network. (2012). *Organ Donation Facts & Information.* Retrieved from Donate Life: www.donatelifeny.org

Parliamantery Office of Science and Technology. (2004). *www.parliament.uk/documents/upload/POSTpn231.pdf.* Retrieved from Organ Transplants: Postnote.

Abadie, A., & Gay, S. (2005). The Impact of Consent Legislation on Cadeveric Organ Donation: A Cross Country Study.

Aderman, D. (1972). Elation, depression and helping behavior. *Journal of Personality and Social Psychology* , 91-101.

AK-DER. (2010, 7). *Toplumsal Cinsiyet, Eşitlik & Başörtüsü Yasaklari.* Retrieved 4 8, 2011, from http://www.ak-der.org/grafik/manager/Dokumanlar/istatistik%20raporu%202010_t%C3%BCrk%C3%A7e.pdf

Alpizar, F., Carlsson, F., & Johansson-Stenman, O. (2008). Anonymity, reciprocity, and conformity: Evidence from voluntary contributions to a national park in Costa Rica. *Journal of Public Economics* , 1047–1060.

Andreoni, J. (1989). Giving with Impure Alturism: Application to charity and ricardian equivalence. *The Journal of Political Economy* , 1447-1458.

Andreoni, J. (1990). Impure Alturism and donations to public goods- a theory of warm glow giving. *Economic Journal* , 464-477.

Bauer, C. (1988). A Direct Mail Customer Purchase Model. *Journal of Direct Marketing* , 16-24.

Becker, G., & Elías, J. (2007). Introducing Incentives in the Market for Live and Cadaveric. *The Journal of Economic Perspectives* , 3-24.

Bekkers, R. (2005). It's Not All in the Ask Effects and Effectiveness of Recruitment Strategies Used by Nonprofits in the Netherlands. *34rd Annual Arnova Conference.* Washington DC, the USA.

Bekkers, R. (2006). Traditional and Health Related Philanthropy: The Role of Resources and Personality. *Social Psychology Quarterly* , 349-366.

Bekkers, R., & Wiepking, P. (2007). Generosity and Philanthropy: A Literature Review. *Working Paper* .

Bennett, R., & Savani, S. (2004). Factors influencing the willingness to donate body parts for transplantation. *Journal of Health and Social Policy* , 61-85.

Bilgel, F. (2010). The impact of presumed consent laws and institutions on deceased organ donation. *The European Journal of Health Economics* , 29-38.

Boulware, L., Ratner, L., Sosa, J., Cooper, L., LaVeist, T., & Powe, N. (2002). Determinants of willingness to donate living related and cadaveric organs: identifying opportunities for intervention. *Transplantation* , 683-691.

Bound, J., Jaeger, D. A., & Baker, R. (1995). “Problems With Instrumental Variables Estimation When the Correlation Between the Instruments andthe Endogenous Explanatory Variables Is Weak. *Journal of the American Statistical Association* , 443–450.

Bryant, W. K., Jeon-Slaughter, H., Kang, H., & Tax, A. (2003). Participation in Philanthropic Activities: Donating Money and Time. *JOURNAL OF CONSUMER POLICY* , 43-73.

Carkoglu, A., & Toprak, B. (2006). *Religion, Society and Politics in Changing Turkey (Değişen Türkiye'de Din, Toplum Siyaset).* Istanbul: TESEV.

Chang, G., Mahanty, H., Ascher, N., & Roberts, J. (2003). Expanding the donor pool: can the Spanish model work in the United States? *American Journal of Transplantation* , 1259–63.

Chick, K. (2010, July 20). *Veil ban: Why Syria joins Europe in barring the niqab*. Retrieved 04 05, 2011, from The Christian Science Monitor: http://www.csmonitor.com/World/Middle-East/2010/0720/Veil-ban-Why-Syria-joins-Europe-in-barring-the-niqab

Choi, J., Laibson, D., & Madrian, C. B. (2004). Plan Design and 401(k) Savings Outcomes. *National Tax Journal* , 275-98.

CIA. (2012). *The World Factbook.* Retrieved 2 7, 2013, from https://www.cia.gov/library/publications/the-world-factbook/geos/sy.html

CIA. (2012). *The World Factbook.* Retrieved 2 7, 2013, from https://www.cia.gov/library/publications/the-world-factbook/geos/bu.html

Cindoglu, D. (2010). *Headscarved Women In Professional Jobs: Revisiting Discrimination In 2010 Turkey* . Istanbul: TESEV.

Clotfelter, C. (1985). *Federal Tax Policy and charitable giving.* Chicago: University of Chicago Press.

Cohen, L. R. (1989). Increasing the supply of transplant organs: The virtues of a futures market. *George Washington Law Review* , 1-51.

Cota, N., Burgess, M., & English, W. (2013). Organ Donation After Circulatory Death. *Anaesthesia Tutorial of The Week 282* .

Crumpler, H., & Grossman, P. (2008). An experimental test of warm glow giving. *An experimental test of warm glow giving* , 1011–1021.

D66. (2012, 08). *Politieke Partij Democraten 66*. Retrieved 11 09, 2012, from D66 wil donorkeuze laten registreren: https://www.d66.nl/actueel/d66-wil-donorkeuze-laten-registreren/

Davis, C. L., & Delmonico, F. L. (2005). Living-Donor Kidney Transplantation: A Review of the Current Practices for the Live Donor. *Journal of American Society of Nephrology* , 2098-2110.

De Neve, K., & Cooper, H. (1999). The Happy Personality: A Meta-Analysis of 137 Personality Traits of Subjective Well-Being. *Psychological Bulletin* , 197-229.

Diamond, W., & Noble, S. (2001). Defensive Responses to Charitable Solicitations. *Journal of Interactive Marketing* , 2-12.

Directorate General on the Status of Women. (2008). Retrieved 10 2011, from http://www.kadininstatusu.gov.tr/upload/mce/eski_site/Pdf/TCEUlusaleylemplani.pdf

Directorate General on the Status of Women. (2011). Retrieved 10 2011, from http://www.kadininstatusu.gov.tr/tr/html/124/Yasal+Haklar+_Genel+Bilgi

Dolan, P., Peasgood, T., & White, M. P. (2008). Do We Really Know What Makes Us Happy? A Review of the Economic Literature on the Factors Associated with Subjective Wellbeing. *Journal of Economic Psychology* , 94-122.

Drago, R., & Williams, C. (2010). *The Gender Wage Gap: 2009.* Washington D.C.: The Institute for Women's Policy Research.

Dunbar, H. (2009, 21 January). Tunisian Women's Current Liberties and Future Plans for Advancement. Illinois, USA.

Dunn, E., Aknin, L., & Norton, M. (2008). Spending money on others promotes happiness. *Science* , 1687-1688.

Ellison, M., McBride, M., Taranto, S., & Delmonico, F. (2002). Living kidney donors in need of kidney transplants: A report from the organ procurement and transplantation network. *Transplantation* , 1349 –1351.

Eurydice. (2008). *The Education System in Turkey, 2007/08.* Brussels: Eurybase: The Education Database on Education Systems in Europe, European Commission.

Falk, A. (2007). Gift exchange in the field. *Econometrica* , 1501-1511.

Field, T., Hernandex-Reif, M., Quintino, O., Schanberg, S., & Kuhn, C. (1998). Elder retired volunteers benefit from giving massage therapy to infants. *Journal of Applied Gerontology* , 229-239.

Filmer, D. (2004). *If You Build It, Will They Come? School availability and school enrollment in 21 poor countries.* Washington, DC:The World Bank: World Bank Policy Research Working Papers (No. 3340).

Filmer, D., & Pritchett, L. (2001). Estimating wealth effects without expenditure data – or tears:an application to educational enrollments in states of India. *Demography* , 115-132.

Forgas, J., Dunn, E., & Granland, S. (2008). Are you being served? An unobtrusive experiment of affective influences on helping in a department store. *European Journal of Social Psychology* , 333-342.

Frank, R. (2004). How not to buy happiness. *Daedalus* , 69-79.

Gevers, S., Janssen, A., & Friele, R. (2004). Consent Systems for Post Mortem Organ Donation in Europe. *European Journal of Health Law* , 175-186.

Gill, M. (2004). Presumed consent, autonomy, and organ donation. *Journal of Medicine and Philosophy* , 37-59.

Glazer, A., & Konrad, K. (1996). A signalling explanation for charity. *American Economic Review* , 1019-1028.

Global Observatory on Donation and Transplantation. (2011, February 9). Retrieved from http://data.transplant-observatory.org/paginas/informes/DatosUsuario.aspx

Gnant, M., Wamser, P., Goetzinger, P., Sautner, T., Steininger, R., & Muehlbacher, F. (1991). The impact of the presumed consent law and a decentralized organ procurement system on organ donation: quadruplication in the number of organ donors. *Transplant Proceedings* , 2685–2686.

Greene, W. (2010). *Fixed Effects Vector Decomposition: A Magical Solution to the Problem of Time Invariant Variables in Fixed Effects Models?* New York.

Hancock, C. (2008). Spatialities of the Secular. *European Journal of Women Studies* , 165-179.

Harbaugh, W. (1998). What do donations buy?: A model of philanthropy based on prestige and warmglow. *Journal of Public Economics* , 269–284.

Harbaugh, W., Myer, U., & Burghart, D. (2007). Neural Respponses to taxation and Voluntary Giving: Real Motives for Charitable Donations. *Science* , 1622-1625.

Hazar Group. (2007). *Türkiye'nin Örtülü Gerçeği (Turkey's Veiled Reality).* Retrieved 10 17, 2011, from http://www.hazargrubu.org/component/content/article/453.html

Healy, K. (2005). The Political Economy of Presumed Consent. *Working Paper* .

HEC, No. 7327 (1982).

HEC, No. 84.15.527 (1984).

Higher Education Council. (2011). *Yuksek Ogretim Kurumu/Universiteler.* Retrieved 10 2011, from http://www.yok.gov.tr/content/view/527/

Hill, M. A., & King, E. M. (1993). Women's Education in Developing Countries: An Overview. In M. A. Hill, & E. M. King, *Women's Education in Developing Countries.* Baltimore, MD: The Johns Hopkins University Press.

Howard, D. H., & Byrne, M. M. (2007). Should We Promote Organ Donor Registries When So Few Registrants Will End Up Being Donors? *Medical Decision Making* , 243-249.

Howard, D. (2007). Producing Organ Donors. *The Journal of Economic Perspectives* , 25-36.

Huck, S., & Rasul, I. (2011). Matched fundraising: Evidence from a natural field experiment. *Journal of Public Economics* , 351–362.

Human Rights Watch. (2009, February). *Discrimination in the Name of Neutrality: Headscarf Bans for Teachers and Civil Servants in Germany.* Retrieved 11 12, 2011, from http://www.hrw.org/sites/default/files/reports/germany0209_web.pdf

ICHEFAP . (2009). *Higher Education Finance and Cost-Sharing in Turkey.*

Imbens, G., & Angrist, J. (1994). Identification and Estimation of Local Average Treatment Effects. *Econometrica* , 467-475.

Isen, A., & Levin, P. (1972). The effect of feeling good on helping: Cookies and kindness. *Journal of Personality and Social Psychology* , 384-388.

Johnson, E. J., Hershey, J., Meszaros, J., & Kunreuther, H. (1993). Framing, Probability Distortions, and Insurance Decisions. *Journal of Risk and Uncertainty* , 35–51.

Johnson, E., & Goldstein, D. (2003). Do Defaults Save Lives? *Science* , 1338-1339.

Jonker, J., Paap, R., & Franses, P. (2000). *Modeling charity donations: target selection, response time and gift size.* Rotterdam: Econometric Institute, Erasmus University.

Judge, H. (2004). The Muslim Headscarf and French Schools. *American Journal of Education* , 1-24.

Kahneman, D., & Deaton, A. (2010). High income improves evaluation of life but not emotional well-being. *Proceedings of the National Academy of Sciences* , 16489-16493.

Kaynakoglu, E., & Toprak, B. (2004). *Women in Labor Market, Senior Management and Politics (İş Yaşamı, Üst Yönetim ve Siyasette Kadın).* Istanbul: TESEV.

Konda. (2007). *Religion, Secularism and Headscarf in Daily Life (Gündelik Yaşamda Din, Laiklik Ve Türban).* Istanbul: Konda.

Landry, C., Lange, J., Price, M., & Rupp, N. (2006). Towards an Understanding of Economics of Charity: Evidence from a Field Experiement. *Quarterly Journal of Economics* , 747-782.

Lindskold, S., Forte, R. A., Haake, C. S., & Schmidt, &. E. (1977). The Effects of Directness of Face-to-Face Requests and Sex of Solicitor on Streetcorner Donations. *The Journal of Social Psychology* , 45-51.

Lyubomirsky, S., Sheldon, K. M., & Schkade, D. (2005). Pursuing happiness: The architecture of sustainable change. *Review of General Psychology* , 111–131.

Madrian, B., & Shea, D. F. (2001). The Power of Suggestion: Inertia in 401(k) Participation and Savings Behavior. *The Quarterly Journal of Economics* , 1149-1187.

Manninen, D., & Evans, R. (1985). Public attitudes and behavior regarding organ donation. *JAMA* , 3111-3115.

Mardell, M. (2006, January 16). *BBC NEWS*. Retrieved 4 5, 2011, from http://news.bbc.co.uk/2/hi/europe/4616664.stm

Matesanz, R., & Miranda, B. (2002). A decade of continuous improvement in cadaveric organ donation: the Spanish model. *Journal of nephrology* , 22-28.

Meer, J., & Rosen, H. (2011). The ABCs of charitable solicitation. *Journal of Public Economics* , 363–371.

Meier, S., & Stutzer, A. (2008). Is volunteering rewarding in itself. *Economica* , 39-59.

Ministry of Education. (2011, 10). *T.C. Millî Eğitim Bakanliği/Özel Öğretim Kurumlari Genel Müdürlüğü.* Retrieved 10 12, 2011, from http://ookgm.meb.gov.tr/kurumlar/dershane.pdf

Mocan, N., & Tekin, E. (2007). The Determinants of the Willingness to Donate an Organ among Young Adults:. *Social Science and Medicine* , 2527-2538.

Moll, J., Krueger, F., Zahn, R., Pardini, M., de Oliveira-Souza, R., & Grafman, J. (2006). Human fronto-mesolimbic networks guide decisions about charitable donation. *Proceedings of the National Academy of Sciences* , 15623-15628.

Mossialos, E., Costa-Font, J., & Rudisill, C. (2008). Does organ donation legislation affect indivudual willingness to donate their own or their relative's organs? Evidence from European Union Survey Data. *BMC Health Services Research* , 8-48.

MSPC. (2011). *Quantitative Information regarding 2011 University Entrance Examination (2011-ÖSYS Yerleştirme Sonuçlarına İlişkin Sayısal Bilgiler).* Retrieved 10 2011, from T.C. Olcme, Secme ve Yerlestirme Merkezi: http://www.osym.gov.tr/dosya/1-58052/h/02-2011-osysyerlestirmesonuclarinailiskinsayisalbilgile-.pdf

National Security Council. (1997, February 28). Decision No: 406 Appendix A.

Newsletter Transplant. (2009). http://www.ont.es/publicaciones/Paginas/Publicaciones.aspx.

Newsletter Transplant. (2010). http://www.ont.es/publicaciones/Paginas/Publicaciones.aspx.

NHS Blood and Transplant. (2012). *Organ Donation Activity Report.*

Official Newspaper. (1981, 11 7). p. 17537.

Official Newspaper. (1982, 10 25). p. 17849.

Official Newspaper. (1982, 11 24). p. 17908.

Official Newspaper. (1987, 01 08). p. 19335.

Official Newspaper. (1988, 12 27). p. 20032.

Official Newspaper. (1989, 12 28). p. 20386.

Official Newspaper. (2006, 07 04). p. 5525.

Onderstal, S., Schram, A., & Soetevent, A. (2011). Door-to-Door Fundraisers Had it Right from the Start.

Orentlicher, D. (2008). Presumed Consent to Organ Donation: Its Rise and Fall in the United States. *Rutgers Law Review* , 61-295.

Oz, M., Kherani, A., Rowe, A., Roels, A., Crandall, C., Tomatis, L., et al. (2003). How to improve organ donation: results of the ISHLT/FACT poll. *Journal of Heart and Lung Transplantation* , 389–410.

Paparde, I. (2010). *Latvian public unwilling to donate organs.* Retrieved 9 29, 2012, from http://ec.europa.eu/health/blood_tissues_organs/docs/art_latvia_2010_en.pdf

Pearson, Y., Bazely, P., Spencer-Plane, T., Chapman, J., & Robertson, P. (1995). A survey of families of brain dead patients: their experiences, attitudes to organ donation and transplantation. *Anaesth Intensive Care* , 88-95.

Ribar, D., & Wilhelm, M. (2002). Altruistic and Joy-of-Giving Motivations in Charitable Behavior. *Journal of Political Economy* , 425-433.

Roels, L., Vanrenterghem, Y., Waer, M., Christiaens, M., Gruwez, J., & Michielsen, P. (1991). 3 Years of Experience with a Presumed Consent Legislation in Belgium - Its Impact on Multiorgan Donation in Comparison with Other European Countries. *Transplantation Proceedings* , 903-904.

Rosenhan, D., Underwood, B., & Moore, B. (1974). Affect moderates self-gratification and alturism. *Journal of Personality and Social Psychology* , 546-552.

Roth, A. (2007). Repugnance as a Constraint on MarketsRepugnance as a Constraint on Markets. *The Journal of Economic Perspectives* , 37-58.

Saharso, S. (2007). Headscarves: A Comparison of Public Thought and Public Policy in Germany and the Netherlands. *Critical Review of International Social and Political Philosophy* , 513-530.

Sargeant, A., Ford, J., & West, D. (2000). Widening the appeal of charity. *International Journal of Nonprofit and Voluntary Sector Marketing* , 318-332.

Schutt, G. (2002). 25 Years of Organ Donation: European Initiatives to increase organ donation. *Transplantation Proceedings* , 2005-2006.

Schuyt, T., Gouwenberg, B., & Bekkers, R. (2011). *Giving in the Netherlands 2011.* Amsterdam: Reed Business.

Siminoff, L., Gordon, N., Hewlett, J., & Arnold, R. (2001). Factors Influencing Families' Consent for Donation of Solid Organs for Transplantation. *Journal of the American Medical Association* , 71-77.

Skjeie, H. (2006). *Gender Equality-Cultural Diversity-Religious Pluralism:Norwegian Case.* Retrieved 4 5, 2011, from http://www.fsw.vu.nl/en/research/research-programmes/sociology/social-change-and-conflict/conference-gender-equality.asp

Staiger, D., & Stock, J. H. (1997). Instrumental Variables Regression with Weak Instruments. *Econometrica* , 557-586.

Stiglitz, J., Sen, A., & Fitoussi, J. (2009). *Report by the Commission on the Measurement of Economic Performance and Social Progress.* OECD.

Tankersley, D., Stowe, C., & Huettel, S. (2007). Altruism is associated with an increased neural response to agency. *Nature Neuroscience* , 150-151.

The White House Office of the Press Secretary. (2012, March 12). *WEEKLY ADDRESS: Women's History Month Highlights the President's Resolve to Pass the Paycheck Fairness Act.* Retrieved from the White House: http://www.whitehouse.gov/the-press-office/2011/03/12/weekly-address-womens-history-month-highlights-presidents-resolve-pass-p

Todd, S., & Lawson, R. (1999). Towards a better understanding of the financial donor: An examination of donor behavior in terms of value structure and demographics. *International Journal of Nonprofit and Voluntary Sector Marketing* , 235-244.

TURKSTAT. (2002). *Educational Statistics/ Educational Expenditures/The amount of education expenditures for service providers by financial sources.* Retrieved 10 2011, from http://tuik.gov.tr/VeriBilgi.do?alt_id=14

TURKSTAT. (2006). *Thoughts on Women's Employment (Family Structure Survey).* Ankara: TURKSTAT.

TURKSTAT. (2002). *TUIK Education Statistics/ Education Expenditures.* Retrieved 10 2011, from http://tuik.gov.tr/VeriBilgi.do?alt_id=14

Vakulenko, A. (2007). Islamic Headscarves and the European Convention on Human Rights. *Social Legal Studies* , 183-199.

Van Diepen, M., Donkers, B., & Franses, P. (2006). *Dynamic and Competitive Effects of Direct Mailings.* Working Paper, Erasmus Institute of Management.

Vohs, K., Mead, N., & Goode, M. (2006). The psychological consequences of money. *Science* , 1154-1156.

Wang, L., & Graddy, E. (2008). Social capital, volunteering and charitable giving. *Voluntas* , 23-42.

Williams, S., & Shiaw, W. (1999). Mood and Organizational Citizenship Behavior: The effects of positive affect on employee organizational citizenship behavior intensions. *Journal of Psychology* , 656-668.

World Economic Forum. (2007). *The Global Gender Gap Report.* Retrieved 4 26, 2013, from http://www3.weforum.org/docs/WEF_GenderGap_Report_2007.pdf

Wunderink, S. (2002). Individual Financial Donations to Charities in the Netherlands: Why, How and How Much? *Journal of Nonprofit & Public Sector Marketing* , 21-39.

Yoruk, B. (2009). How responsive are Charitable Donors to Requests to Give. *Journal of Public Economics* , 1111-1117.